

UNIVERSITY OF NAPLES FEDERICO II



DOCTORAL SCHOOL IN CIVIL ENGINEERING

PHD PROGRAM IN HYDRAULIC SYSTEMS,
TRANSPORTATION AND URBAN DEVELOPMENT
ENGINEERING

DOCTORAL THESIS

*Modelling Tools and Simulation Platforms for Advanced Driver
Assistance Systems*

Maria Russo Spena

Tutor:

Prof. Eng. Gennaro Nicola Bifulco

Ph.D. Coordinator:

Prof. Arch. Elvira Petroncelli

Examiner:

Prof. Eng. Luca D'Acierno

*'Never argue with an idiot. It
drags you to his level and
beat you with the experience'
(cit. O. Wilde)*

Overview

Recent developments in Advanced Driver Assistance Systems (ADAS) lead to the concept of *eco-driving* support systems, which assist drivers in controlling vehicles in a sustainable way by reducing fuel consumption and carbon dioxide emissions. In this thesis, an algorithm for an ecological Adaptive Cruise Control (*ecoACC*) system is designated. In order to test the system in MIL (Model in the Loop) and DIL (Driver In the Loop) environments, an important task is the specification of a fuel consumption model.

In the first part of the thesis an overview is presented about fuel consumption models and several eco-devices developed by different authors. In the second part an experimental campaign (with a sample of 100 drivers) is carried out in order to specify and to validate the model. The survey is supported by the DRIVEIN² (DRIVER monitoring: technologies, methodologies, and IN-Vehicle Innovative systems) project. As the major data are collected using the OBD-II port, and in particular the data related the instantaneous fuel consumption, a validation of the former data is carried out in collaboration with the CNR (*Consiglio Nazionale delle Ricerche – Istituto motori*).

To carry out a more realistic simulation a vehicle dynamics model was coupled with the ADAS model. The purpose of the implemented vehicle dynamics model is to simulate, in real time, the vehicle's dynamic behaviour, thanks to continuous integration of the balance equations regulating the longitudinal and lateral vehicle motion. Furthermore, a model able to provide an estimation of real-time fuel consumption of the vehicle is developed: such a model is embedded in the vehicle model discussed in the following, in order to optimize consumption performance during driving task.

Index of Figures

Figure 1-Total model diagram proposed by Van der Voort et al. (2001)	11
Figure 2 – CompactDAQ and pc controller	15
Figure 3 – TRW Autocruise AC20 radar mounted behind, Xsense MTi-G	16
Figure 4 – Front and rear cameras	17
Figure 5 – OBD II port locates under the dashboard	18
Figure 6 –Instrumented Vehicle equipped with PEMS	19
Figure 7 – Diagram of power supply system of SEMTECH	20
Figure 8 – Scheme of devices mounted during the on-road test.....	21
Figure 9 – DICEA Driving simulator	21
Figure 10 – The experimental path	26
Figure 11 – The experimental path	29
Figure 12 – Percentage differences between aggregate data collected by the OBD port and by the PEMS for each driver.	39
Figure 13 – Empirical cumulative mass functions (ecmf) of gap between data collected by the OBD port and by the PEMS for Driver 6 (on the top) and Driver 3 (on the bottom) using different time slots.	41
Figure 14 – Empirical probability mass function (epmf) for aggregate data each 3 minutes differentiating for the median value of RPM	42
Figure 15 - Dependence of fuel metering on: (lefts side) Gas Pedal; (right side) Intake Air	45
Figure 16- Screenshot of the results for the null model (36)	46
Figure 17 – Box whisker plot of Fuel metering vs subjects	47
Figure 18 – Histogram of marginal residuals	54
Figure 19 – qq plot conditional residuals vs fitted values	54

Figure 20 – Empirical probability mass function (epmf) and empirical cumulative mass function (ecmf) vs the percentage error (ERR_{inst}) of the two sub-samples.....	57
Figure 21 – Empirical probability mass function (epmf) and empirical cumulative mass function (ecmf) of the aggregate error (ERR_{aggr}) for the calibration and the control sub-samples.	59
Figure 22 – Scheme of eco-drive system	61
Figure 23 – Architecture implemented in Matlab/Simulink.....	62
Figure 24 – ‘Simulated Driver’ sub-system.....	64
Figure 25 – Screenshot showing the vehicle dynamics Simulink model.....	66
Figure 26 – “Input/Output” Simulink subsystem.....	67
Figure 27 – “Control Unit Simulink subsystem”	67
Figure 28 – “Tires” Simulink subsystem	68
Figure 29 - “Vehicle Dynamics” Simulink subsystem	69
Figure 30 – “Gearbox Control Unit” Simulink subsystem	69
Figure 31 – Architecture implemented in DIL environment	71
Figure 32 – Simulink/SCANeR communication protocol	72
Figure 33 – SCANeR Studio window.....	73
Figure 34 - SCANeR Studio main	75
Figure 35 – (on the top) Leaves used for driver feedback, (on the bottom) Screenshot of graphical interface provides to driver.....	76

Index of Table

Table 1 – The sample	24
Table 2 - Correlation matrix	45
Table 3- Parameters of lme model (41)	48
Table 4 – Standardized beta for model (41).....	49
Table 5 – Parameters of lme model (42).....	50
Table 6 – Standardized beta for model (42).....	50
Table 7 - LogLikelihood Ratio test for lme models (41) - (42)	51
Table 8 – Variance-Covariance matrix of random effects for lme model (42)..	51
Table 9 - LogLikelihood Ratio test for lme models (43) - (42)	52
Table 10 – Parameters of lme model (46).....	55
Table 11 - Variance-Covariance matrix of random effects for lme model (46)	56
Table 12 – RMSE and MAPE computed for the two sub-samples.....	58
Table 13 – RMSE and MAPE for aggregate measurements.....	59

Table of contents

Overview	I
Index of Figures	II
Index of Table	IV
Table of contents	V
Introduction	1
Chapter I – State of the art	4
1.1 Beyond the state of the art	13
Chapter II - Strategies and tools	14
Introduction	14
2.1 Tools	14
2.1.1 Instrumented Vehicle	14
2.1.2 Portable Emission Measurement System	18
2.1.3 Twins Driving Simulators	21
2.2 Experimental surveys and data collection	23
2.2.1 DriveIn ² experiment	23
2.2.2 DriveIn ² experiment add-on	28
Chapter III – Fuel consumption model	30
Introduction	30
3.1 Linear mixed effect model	30
3.1.1 Specification for all data	33
3.1.2 Maximum-Likelihood Estimation	34
3.2 Validation of OBD fuel consumption	37
3.3 Specification of fuel consumption model	42
Chapter IV – Implementation of ADAS in MIL and DIL environment	60
Introduction	60
4.1 The implemented model in MIL	60
4.2 The implemented model in DIL	70
Conclusion	77
References	79

Introduction

The research here presented takes place in the frame of driving behaviour studies, and the design of Advanced Driving Assistance Systems (ADASs). ADASs are developed using electronic control units (ECUs) that are aimed to supervise the safety, the comfort and/or the efficiency of the driving.

Their design is carried out by using the so-called automotive V-Cycle, a process of analysis, development and prototyping that envisages a first phase of conception of the control logic characteristics and mechanisms, a second phase in which the software needed to manage such a logic is developed, and a third phase consisting on the implementation of the ECU and the evaluation of its interaction with the other components (hardware sensors, actuators and other devices, control units) of the vehicle.

The first stage is called Model In the Loop (MIL), the model of the control logic interacts with the models (already certified and tested) that represent the operation of the entire vehicle (as well as the environment and traffic).

In the second phase, called Software In the Loop (SIL), the controller management software that implements the control logic interacts with other software codes and other system component models in a simulated

environment; however, the input and output of the tested software are of the same kind and type of real ones.

In the third stage, known as Hardware In the Loop (HIL), a prototype controller is actually developed (by using embedded programming techniques), and the obtained hardware-component is plugged into a mixed testing environment composed by both hardware and other software components (already developed and certified). The actual hardware/software mix depends on the particular aim of the experiment and is aimed at fully reconstructing the vehicle architecture.

The research here described falls in the MIL context and is aimed at enhancing existing ADAS testing platforms (generally developed in Matlab/Simulink) with the possibility of including more sophisticated models that takes into account the interaction of the controlled vehicle with other vehicles (traffic), and with the drivers' behaviours. Particular focus is dedicated, in the previous framework, to the development of a fuel consumption module. It will be verified the possibility of including at the MIL level also observation of drivers' reactions in a virtual environment (using Driving Simulators), thus creating a DIL (Driver In the Loop) level.

The choice of ADAS aimed to reduce fuel consumption and greenhouses gas emission, adopting an eco-driving style, is due to the fact that the fuel saving is a sensible topic. In fact, on 24 June 2013, the Environment Committee of the European Parliament approved an agreement on new rules to reduce the CO₂ emissions of passenger cars. According to the text by 2021, for new vehicles sold in the EU, CO₂ emissions must be reduced by 28% (from the current 132 g/Km to 95 g/Km). The text was approved by the plenary in Strasbourg on 24-27 February 2014. A possible approach to the problem is to act at the vehicle level, understanding what mostly affects fuel consumption, and developing the design of Advanced Driver Assistance Systems (ADASs) that help the drivers in adopting an eco-driving style.

The thesis is organized as follows: section 1 resumes the literature with respect to the various approaches used for the analysis of fuel consumption and

different eco-drive tools implemented and tested in real and virtual environment; in section 2 the two experiments, the tools adopted and the collected data are presented; in section 3 validation of OBD fuel data and characterization of fuel model are discussed, in section 4, implementation of eco-drive architecture in MIL and DIL are presented; and finally, some conclusions are drawn.

Chapter I – State of the art

Transport systems are affected by serious negative externalities in terms of both pollution (Davitashvili, 2009) and noise (Pamanikabud, 2009). Indeed it has been estimated that the impact of the transport sector is in the range of 20%–40% in terms of consumption of fossil fuels and emission of greenhouse gases and particulate (Carteni, 2014). To manage and reduce these negative impacts, many fuel-consumption and air-pollution control policies have been adopted, such as the introduction of carbon-based vehicle tax systems (Chi, Guo, Zheng, & Zhang, 2014), incentives for the adoption of hybrid electric vehicle (Zhang, Xie, Rao, & Liang, 2014), improvements in vehicle technologies (catalytic converter, alternative fuels, engine efficiency), congestion charging, and incentives for using public transport. Furthermore it's also important to pay attention on the effect of traffic calming measures on the environmental and energy impacts. These involve strategic measures to improve safety and enhance quality of life by reducing vehicle speed and/or cut-through volumes on streets or street networks. In fact as shown in the case study presented by Ahn et al. (2009), while traffic-calming measures reduce vehicle speed on neighbourhood streets enhancing road safety, these measures can result in significantly higher fuel consumption and emission rates when drivers accelerate aggressively.

A different effective and promising technique to increase fuel saving is eco-driving, which is essentially based on observing and controlling the driver behaviour. Eco-driving is defined as a decision-making process, which will influence fuel-savings and emissions intensity of a vehicle to reduce its environmental impact (Sivak & Schoettle, 2012).

At the transportation network level, eco-driving acts by implementing proper instruments for the evaluation and implementation of strategies for the reduction of traffic congestion [(D'Acierno, Montella, & Gallo, 2011), (Gori, La Spada, Mannini, & Nigro, 2014)] and for optimum route choice in order to

limit CO₂ emission and fuel consumption (eco-route, (Sivak & Schoettle, 2012)). Eco-driving at the single vehicle/driver level can be achieved through an understanding what primarily affects fuel consumption, and by developing ADASs that help drivers in adopting efficient driving styles in terms of emissions and energy consumption. At both levels, the effects of the adopted solutions can be investigated by using microscopic simulation tools, which allow the simulation of the interaction of vehicles in a traffic stream and the resulting kinematics (and dynamics) of the vehicles. Microscopic models make it possible to explicitly take into account different driving behaviours (Mauro, Giuffrè, Grana, & Chiappone, 2014) and engine characteristics [(Donateo, 2013), (De Luca, Di Pace, & Marano, 2015)].

In the case of ADAS oriented to eco-driving, a model able to predict fuel consumption is a pre-requisite. Fuel consumption models can be implemented by following three different approaches: macro models, meso models and micro models. Macro models are able to predict fuel consumption for a relatively large region and for rather long periods of time; of course, for the purposes of this research they will not be processed. Meso models are based on average parameters (e.g., the average speed of the vehicle), and can be divided into two major categories: models based on correction factors [(Ntziachristos, Gkatzoflias, Kouridis, & Samaras, 2009), (Brzezinski & Newell, 1998)], and models based on vehicle specific power (VSP) which is the power exerted per unit mass [(Koupal, Cumberworth, Michaels, Beardsley, & Brzezinski), (Hui, Zhang, Yao, & Wang, 2007)]. Consumption calculated using these approaches reflects the average consumption for a certain class of vehicles, often presenting some deviance in results when considering a specific vehicle (Zhou, Huang, Weifeng, & Li, 2013).

This thesis focuses on the microscopic approach, in which developed models can be classified according to different criteria. First and foremost is the fuel supply—gasoline, diesel and newly developed hybrid vehicles. Second is the basis of the models. Some of them have as a basic principle the traction-law. These models estimate the fuel consumption as a function of the mass of the vehicle, the aerodynamic drag coefficient, the vehicle frontal area, the

vehicle acceleration and speed, the road gradient, etcetera. In particular, Post et al. (1984) developed a fuel consumption model that is a linear function of the instantaneous power demand; the model was built starting from the available data of chassis dynamometer experiments of 177 in-use vehicle. This model was subsequently enhanced in later publications [(Akcelik, 1989), (Leung & Williams, 2000)]. In other cases, fuel consumption has been indirectly inferred from exhaust gas emissions, even in real time, using the carbon balance method (e.g., Rakha et al. (2004)). An important category concerns models based on laboratory-tests performed on a chassis dynamometer, that implements standard driving cycles such as the European NEDC (New European Driving Cycle), and the U.S. FTP (Federal Test Procedure) [(Wong, 2001), (Rakha, Ahn, Moran, Searens, & Van den Bulck, 2011)].

For the scope of this thesis, it is worth pointing out approaches that make it possible to compute instantaneous fuel consumption, possibly at real-time. Moreover, these models are required for vehicle-specific application, thus it can be applied and specified for a each given vehicle it is applied to, without the need to be representative of the dispersion among several vehicles. For instance, Ahn (1998) developed a nonlinear regression model, in order to predict fuel consumption, as a polynomial combination of the instantaneous speed and acceleration, using different regression coefficients in the deceleration or acceleration phase, as shown in equation (1), in order to construct a dual-regime model. The model was calibrated using data collected at the ORNL (Oak Ridge National Laboratory) using a chassis dynamometer, data gathered using an on-board emission measurement device, and also data gathered by EPA (Environmental Protection Agency). This model differs from the others because, while these last were developed as power-demand models, the VT-Micro model was developed as a statistical model.

$$F(t) = \begin{cases} \exp\left(\sum_{i=1}^3 \sum_{j=1}^3 L_{ij} v(t)^i a(t)^j\right) & a(t) \geq 0 \\ \exp\left(\sum_{i=1}^3 \sum_{j=1}^3 M_{ij} v(t)^i a(t)^j\right) & a(t) < 0 \end{cases} \quad (1)$$

Where L_{ij} and M_{ij} are model parameters at speed exponent i and acceleration j , $v(t)$ is the instantaneous vehicle speed at instant t expressed in kilometers per hour; and $a(t)$ is the instantaneous vehicle acceleration (km/h/s).

In the same period, the University of California (Barth, et al., 2000) presented the Comprehensive Modal Emissions Model (CMEM) which is able to predict instantaneous fuel consumption and emissions in three different driving cycles for both light and heavy vehicles; the model uses as input variables the vehicle kinematic (acceleration, velocity, etc.), the characteristics of the road (e.g., slope), and motor data (such as the coefficient of cold starting, and the coefficient of engine friction), as shown in (2) - (5) :

$$FR \left[\frac{g}{s} \right] = \left(K \cdot N \cdot V + \frac{P}{\eta} \right) \frac{1}{43.2} (1 + b_1 (N - N_0)^2) \quad (2)$$

Where FR is fuel rate grams/second, P is engine power output in kW (eq. (4a)), K is the engine friction factor (described in eq. (3)), N is engine speed (revolutions per second), N_0 is the idling engine speed (eq. (5)), V is engine displacement (liter), and $\eta \approx 0.45$ is a measure of indicated efficiency for diesel engine. $C \approx 0.00125$ and $b_1 \approx 10^{-4}$ are coefficients; 43.2 kJ/g is the lower heating value of a typical diesel fuel. In particular:

$$K = K_0 (1 + C(N - N_0)) \quad (3)$$

$$P[kW] = \frac{P_{tract}}{\varepsilon} + P_{acc} \quad (4a)$$

$$P_{tract}[kW] = A v + B v + C v + M (0.447a + g \sin(\theta))v \frac{0.447}{1000} \quad (4b)$$

$$N_0 = 30 \sqrt{\frac{3.0}{V}} \quad (5)$$

Where K_0 is constant engine friction factor that varies depending on the vehicle. P_{tract} e P_{acc} are respectively the total tractive power requirements and the engine power demand associated with the operation of vehicle accessories such as air conditioning usage; ε is vehicle drivetrain efficiency. In particular the total tractive power requirements, P_{tract} , placed on the vehicle (at the wheels) is described in eq. (4b), where M is the vehicle mass with appropriate inertial correction for rotating and reciprocating parts (kg), v is speed (mph), a is acceleration (mph/s²), g is the gravitational constant (9.81 m/s²), and θ is the road grade angle in degrees. The coefficients A , B and C involve rolling resistance, speed-correction t rolling resistance, and air drag factors.

Moreover all models presented in the literature produce a bang-bang type of control system (i.e. optimal control entails full-throttle input and full braking input); to overcome this type of problem, Rakha et al. (2011) developed the Virginia Tech Comprehensive Power- based Fuel consumption Model (VT-CPFM). In fact a model that results in a bang-bang control system would indicate that the optimum fuel economy control would be to accelerate at full throttle in order to reduce the acceleration time. This type of control, which is obviously incorrect, would induce the driver to drive as aggressively as possible in order to minimize her/his fuel savings. This is due to the fact that the partial derivative of fuel consumption rate (F) with respect to the engine torque (T) isn't a function of engine torque, or $\delta F / \delta T \neq f(T)$ (Saerens, Diehl, & Van den Bulck, 2010). As shown below eq. (6) the model is a second order polynomial as a function of the instantaneous vehicle power $P(t)$:

$$F(t) = \begin{cases} \alpha_0 + \alpha_1 P(t) + \alpha_2 P(t)^2 & P(t) \geq 0 \\ \alpha_0 & P(t) < 0 \end{cases} \quad (6)$$

Here α_0 , α_1 and α_2 are the model parameters and $P(t)$ is the instantaneous total power exerted by the vehicle at any instant t (kw), computed as:

$$P(t) = \left(\frac{R(t) + 1.04 m a(t)}{3600 \eta_d} \right) v(t) \quad (7)$$

With m corresponds to vehicle mass (kg), $a(t)$ and $v(t)$ are respectively the instantaneous acceleration and velocity at instant t , η_d is the driveline efficiency and $R(t)$ is the resistance force given by the following equation:

$$R(t) = \frac{\rho}{25.91} C_d C_h A_f v^2(t) + mg \frac{C_{r0}}{1000} (c_{r1} v(t) + c_{r2}) + mg G(t) \quad (8)$$

Equation (8) computes the sum of the aerodynamic, rolling and grade resistance forces acting on the vehicle. The parameters presented in that equations are: air density a level sea and a temperature of 15°C (ρ), vehicle drag coefficient (C_d), altitude correction factor (C_h), vehicle frontal area (A_f), rolling resistance constants (C_{r0} , C_{r1} , C_{r2}), total vehicle mass (m) and at last roadway grade at instant t ($G(t)$).

Finally, in (Lee, Park, Jung, & Yoo, 2011), using data recorded by the OBD port, a nonlinear regression model for a gasoline vehicle with an automatic gearbox was developed, utilizing as experimental context 5 km

driving route of urban area. In particular three different specifications were tested: the first one expressed the fuel consumption as a quadratic function of RPM, the second as a linear function of throttle and the last as a linear polynomial model of RPM and throttle. It is worth noting the respective determination coefficients (R^2): 0.76, 0.81 and 0.71. In particular:

$$FUEL_{RPM} = a x^2 + bx + c \quad (9)$$

$$FUEL_{TPS} = d y + e \quad (10)$$

$$FUEL_{poly22} = p_{00} + p_{10}x + p_{01}y + p_{20}x^2 + p_{11}xy + p_{02}y^2 \quad (11)$$

Here x represented the instantaneous engine RPM (Revolution Per Minute) measurement, while y is the TPS (Throttle Position Sensor) measurement.

Concerning the eco-driving advisory, the University of Twente (Netherlands) in collaboration with the School of Transportation and Society (Sweden) have developed a fuel-efficiency support tool capable of performing real-time control of consumption and providing both positive and negative video feedback while driving. The support tool includes the so-called normative model, which back-calculates the minimal fuel consumption for manoeuvres carried out (Van der Voort, Dougherty, & van Maarseveen, 2001). As shown in the figure below (Figure 1), this advisory is composed by 5 different stages.

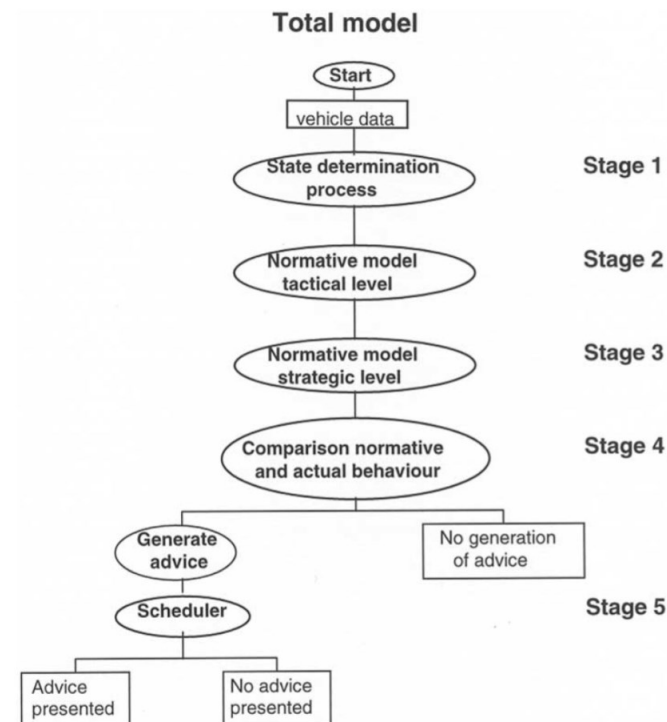


Figure 1-Total model diagram proposed by Van der Voort et al. (2001)

In particular it consists of a module able to determine the current state of vehicle (state determination process– stage 1) such as cruising, decelerating, stationary and so on. A normative model divided into two parts: the tactical level (stage 2) concerned only with the immediate past, in particular it concerns the last time a state change took place. A series of identical states is therefore grouped together and defined as a manoeuvre. Than the second one, the strategic level consists of a set of rules, which have been derived from the studies concerning optimum fuel consumption (stage 3). The next step of the data processing module is the generation of advice to the driver. This is done by comparing actual behaviour against normative optimal behaviour (stage 4). Finally before the advice is given to the driver the scheduler sub-system (stage 5) ensures that there aren't any conflicts between the tactical and strategic level, which could prove confusing to the driver and finally checking the advice doesn't create dangerous situation in the current traffic context. This fuel-efficiency support tool is tested through a driving simulator experiment carried out at the TNO Human Factors Research Institute in The Netherlands. The experiment involved 88 male subjects, and it revealed that drivers are able to reduce overall fuel consumption by 16% comparing with “normal driving”, in

particular there is a reduction of 9% of fuel consumption due to the fuel efficiently driving style adopted by the participants, thus, the tool gave an additional reduction of 7%.

Similarly, Larsson et al. (2009) developed an Acceleration Advisor (AA) able to signal the driver, causing a resistance on the gas pedal if s/he is accelerating too hardly. The values of the speed of pedal depression and the initial resistance were chosen after comparing driving patterns from fixed test runs involving 16 combinations of resistance variables being systematically changed. Driving pattern parameters concerning fuel efficiency (e.g., percentage time of high acceleration) were compared for each test run, as well as the perceived acceptance of the resistance level based on relative time consumption. In a test carried out in three different postal districts located in Southern Sweden, the tool was installed in four postal delivery vehicles between September and December 2005. In particular the vehicles were driven for six weeks without the AA and the other six weeks with the AA activated. The study shows that the AA had a positive effect on emissions, on two of the three routes, but no significant reduction in fuel consumption was observed. The positive note is that the strong acceleration was significantly reduced that indicates the drivers had complied with the advisor.

In fact another important aspect that concerns the ADAS world is related to the kind of feedback to give to a driver. In fact for what concerns the automotive industry, the human is cut-off from the development of an ADAS system, with the high probability that a driver could turn off the system if s/he considered that the adviser tool is too annoying. On the other hand, for the research institute the human acceptability is a crucial task in the development of ADAS system, not only for what concerns the realistic behaviour of the system in the real traffic context but also for the typology of advice. In fact about the eco-driving assistance system interface, Jamson et al. (2015) investigated and compared 12 different potential eco-driving interfaces using a driving simulator. Different visual, visual-auditory and haptic interface systems are tested in a short section of road; demonstrating that it is recommended to consider the strong force feedback system (haptic system), but the authors recommend to test this solution during a longer drive; indeed, for what concerns

visual-auditory systems potential annoyance from prolonged use need to be considered. Finally, about visual systems, the authors recommended to use advices able to indicate how the driver can improve his/her behaviour in respect of fuel efficiency reduction.

1.1 *Beyond the state of the art*

Here a model with a two-fold aim is proposed. First of all this model is intended to be integrated into a larger platform in which it participates to the simulation as a whole by estimating the fuel-consumption, as well as its variations in dependence on the eco-drive logic applied by the developed ADAS. As a secondary aim, the model can be adopted once the ADAS is in place, because, by identifying in real-time the actual value of its parameters, it is possible to profile driving behaviour with respect to eco-efficiency. Both of the previous aims determine some requirements to the proposed fuel-consumption model. Indeed, the model has to be as simple as possible in order to be easily employed and identified in real-time. As well, it must be based on simple variables that can be easily obtained in (simple) ADAS design and testing platforms, and by low-cost aftermarket or OEM (On-road Emission Measurement) solutions installed on vehicles. Finally, the implemented system has to be a “next-generation” eco-drive system, in the sense that it is able to observe the driver's behaviour and the current fuel consumption, as well as to simulate the fuel consumption in the same traffic conditions and stimuli the driver is subject to, and finally to compare the actual fuel consumption with those of an *eco-driver*. In this way the system gives a reward to the driver if his/her fuel consumption is lower than those of the eco driver, vice versa the system indicates the possibility to improve his/her driving style.

Chapter II - Strategies and tools

Introduction

In order to develop a fuel consumption model and validate the fuel data using the OBD port available on vehicles, it's necessary to carry out two kind of experimental campaigns: the first one in order to collect data for fuel consumption estimation, the second one in order to evaluate the reliability of fuel data using OBD port. In the former case an instrumented vehicle is used. Indeed in the latter case, a portable emission measurement system (PEMS) was employed, mounted on the same Instrumented Vehicle. Finally, in order to develop and evaluate the driver acceptability of the *eco-drive* system a driving simulator has been employed. In this chapter the strategies (two experimental campaigns) are discussed, as well as the tools (instrumented vehicle, PEMS and driving simulator) used in order to develop the *eco-drive* system.

2.1 Tools

In this paragraph the tools used for the thesis purpose will be illustrated. In particular the Instrumented Vehicle, the Portable Emission Measurement System (PEMS) and finally a static Driving Simulator are discussed in detail.

2.1.1 Instrumented Vehicle

The Instrumented Vehicle (IV) used in the research is a Fiat Multipla owned by the Department of Civil, Architectural and Environmental Engineering (D.I.C.E.A. - Dipartimento di Ingegneria Civile, Edile ed Ambientale) of the University of Naples "Federico II". Thanks to the DRIVE IN2 project the IV has been upgraded. Some of the upgrades were also suggested by the study reported in (Bifulco, Pariota, Simonelli, & Di Pace, 2011), where the advantages of equipping an IV with different sets of instruments, each with a different level of accuracy, are discussed, as well as

the general level of accuracy reached by the system as a whole by using information fusion techniques.

As described in Bifulco et al. (2012) the core of the system is an National Instruments CompactDAQ chassis, a modular data acquisition system for gigabit ethernet with eight measurement-specific modules. CompactDAQ is the data acquisition device for all the sensors in the vehicle, except the video cameras.



Figure 2 – CompactDAQ and pc controller

It receives digital and analog signals, as well as data and communications via the CAN. It is connected to a PC controller with an Intel Pentium I7 six-core processor. It has 32 GB of RAM, two 500 GB SSDs and three 1TB HDDs. The synchronization and acquisition software is based on LabVIEW. The on-board power supply is partly in direct current (from 5V to 24V depending on the device) and partly in 220V AC. Alternating current is supplied by a 500W inverter. The energy is supplied by two 12V (100Ah) service batteries able to provide safe, uninterrupted power for up to 5 hours. In order to allow for longer experimental sessions, the batteries are recharged while the vehicle runs by means of a Sterling Power active power split charger. A Topcon Global Positioning System (GPS) with a Differential GPS antenna supplies positioning and kinematic data; in this task it is integrated with a dead- reckoning technique based on an Xsense MTi-G Accelerometer (Figure 3, on the right). This is a MEMS Inertial Measurement Unit (IMU), in turn aided by another internal GPS, it is located along the transmission tunnel between the front driver's and

passenger's seats. Two TRW Autocruise AC20 radars are mounted, hidden behind the bumper, on the front (Figure 3, on the left) and on the back of the vehicle to enable collection of headway data (relative distance and speed of up to five surrounding vehicles for each radar).



Figure 3 – TRW Autocruise AC20 radar mounted behind, Xsense MTi-G

Three potentiometers are used to collect data on the position of the brake pedal, the gas pedal and clutch pedal. The steering wheel angle is measured by a draw-wire displacement sensor and the angular velocity by a gyroscope (in addition to the Xsense IMU). Video data are collected by four Basler industrial cameras: 170 degree color giga-ethernet acA1300-30. The cameras have a resolution of 1294 x 964 pixels and a frame rate of 30 fps. Two cameras are oriented forward and backward outside the vehicle. One camera points toward the driver inside the vehicle. The last camera is positioned so that the lateral lane position of the vehicle can be detected. The front and rear cameras record the behaviour of the surrounding traffic and the characteristics of the road environment (Figure 4). The presence of video cameras accounts for the large amount of storage space in the control PC. Video and data acquisition (from all sensors) are synchronized by an external trigger signal.



Figure 4 – Front and rear cameras

The equipment is completed by a touch-screen panel PC allowing data acquisition to be managed and monitored and workload tests to be performed while driving.

Finally, for the aim of this thesis the OBD-II port signals are also used, and these data are synchronized with the others. OBD-II data port is located near the driver, usually under the dashboard as shown in Figure 5.

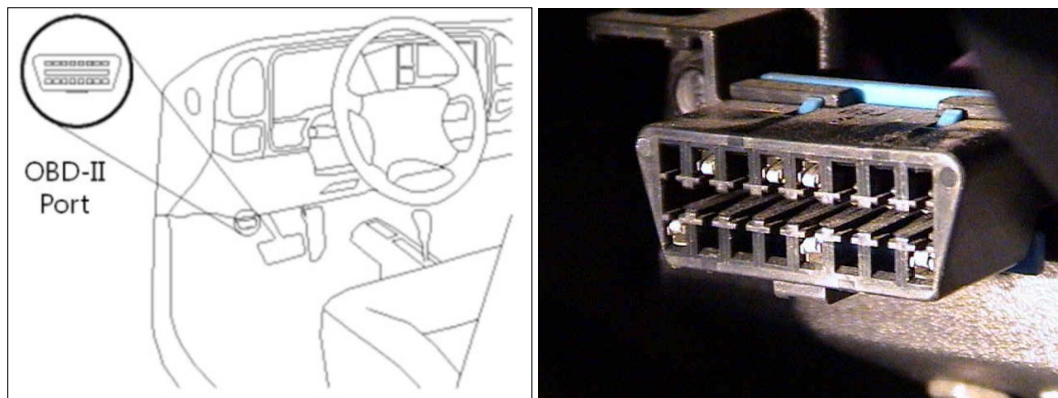


Figure 5 – OBD II port locates under the dashboard

OBD-II provides access to numerous data from the engine control unit (ECU) and offers a valuable source of information when troubleshooting problems inside a vehicle. The SAE J1979 standard defines a method for requesting various diagnostic data and a list of standard parameters that might be available from the ECU. In our case, using the OBD-II port the following data are collected: odometer measures, velocity, rpm, the excursion of gas pedal, temperature, the instantaneous fuel consumption and intake air into the combustion chamber, air temperature, battery voltage, EGR, and finally dummy variables of Air Conditioner (on/off), brake and clutch pedal (on/off). These variables are update and, then collected, with a frequency of 10 Hz.

2.1.2 Portable Emission Measurement System

In order to validate the OBD fuel data, a second experiment has been carried out using a PEMS (Portable Emission Measurement System) in collaboration with the CNR - Istituto Motori of Naples. Portable emissions measurement systems are complete sets of emission measurement instruments that can be carried on board of the vehicle under study. Such systems can provide instantaneous emission rates of selected pollutants with satisfactory levels of accuracy. PEMS typically measure instantaneous raw exhaust emissions of CO₂, CO, NO_x and THC. The main advantage of on-board methods is that they can provide long series of emissions data for a particular vehicle driven under a wide range of operational/duty cycles and driving and ambient conditions (Cicero-Fernández, Long, & Winer, 1997), including some that would otherwise be difficult to replicate in the laboratory (e.g., large road

gradients). The installation of PEMS on several vehicles of various categories can lead to a large database of emission values from vehicles of different technologies driven under different driving and environmental conditions. The main limitations of PEMS include the reduced range of measurable pollutants, the added mass (of approximately 30 to 70kg) that may bias the measurement (especially for light-weight cars), and the reduced reproducibility due to real-world sources of variability. Also, the range of pollutants that can be measured with PEMS is still limited in comparison to laboratory measurements (Franco García, 2014).

The device owned by the Istituto Motori is a Semtech-Sensor, mounted in the rear of the IV, as shown in the figure below (Figure 6). PEMS can record instantaneous pollutant concentration and exhaust flow data streams at 1 Hz.

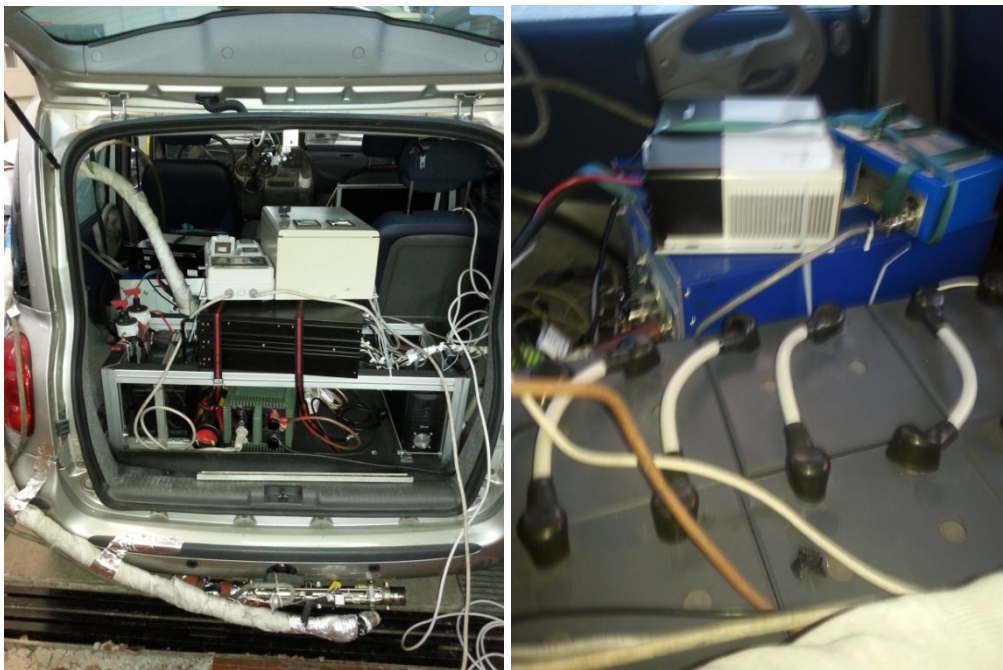


Figure 6 –Instrumented Vehicle equipped with PEMS

The Semtech is powered by six batteries, each one at 4 V for a total of 24 V. In the Figure 7 is reported the diagram of power supply system, as shown there is a connection in series among the six battery, in order to obtain a voltage of 24V. Moreover this group of batteries is connected to a switchboard, which has two safety switches (one on the positive pole, and the other one on the negative pole).

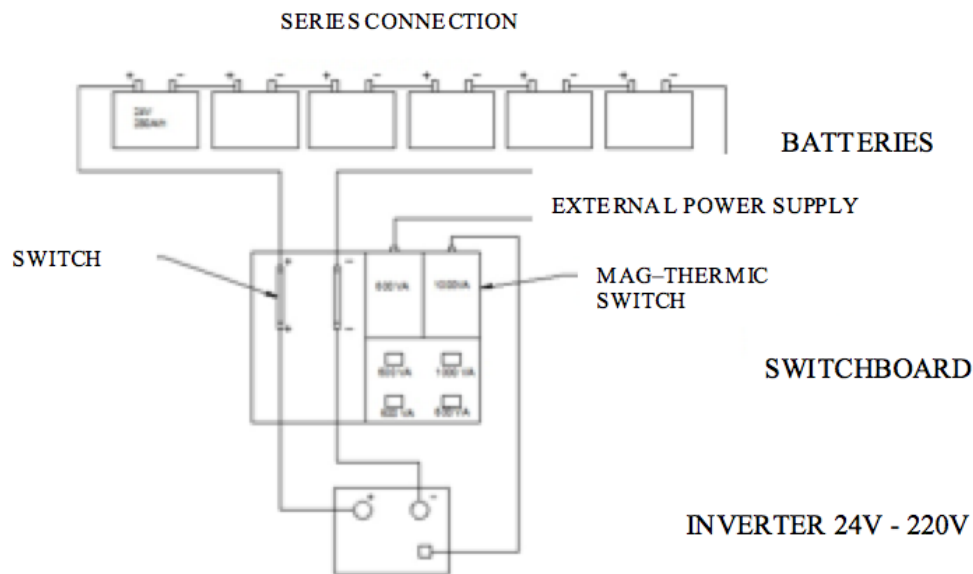


Figure 7 – Diagram of power supply system of SEMTECH

These two switches also connect the batteries with an inverter; in this way the input tension of 24v is transformed in AC (alternating current) 220v. In this way, it's possible to supply jacks of switchboard with a nominal power of 1000 VA (Volt-Ampere). The Semtech, and all the devices necessary to carry out the road test, are connected to these jacks (Figure 7).

Finally, as illustrated in Figure 8 below, several devices have been mounted on the IV in order to evaluate and record the environment conditions and tailpipe exhaust gas:

- Semtech for the analysis of tailpipe exhaust gas,
- Exhaust Flow Meter measurements (EFM),
- GPS,
- Notebook.

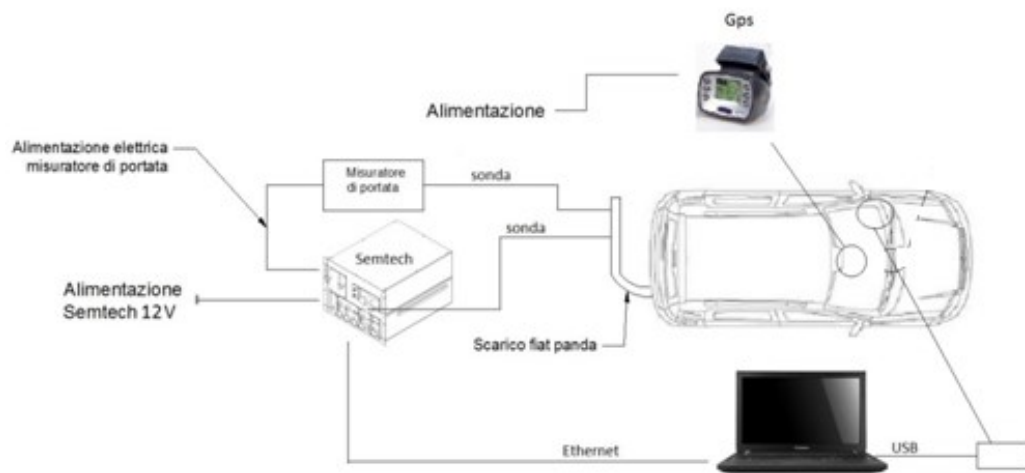


Figure 8 – Scheme of devices mounted during the on-road test

2.1.3 Twins Driving Simulators

The two-twins static driving simulators (DSs), owned by the Dipartimento di Ingegneria Civile Edile ed Ambientale (DICEA) were assembled in May 2012 as part of the DRIVE IN² project. Both the DSs have the same hardware and software in order to enable a complete interaction in the same virtual driving scenario. Each of the cockpits consists on: mono-seat, steering wheel, pedals and gear lever.



Figure 9 – DICEA Driving simulator

The dashboard, the various indicators (as speedometer, fuel level, gear engaged, etc.), and the side/rear-view mirrors are reproduced virtually. The gearshift can be both manual and automatic. All control interfaces provide a realistic response. As an example, the pedals are provided with a dynamic system of force feedback and the torque at the wheel is reproduced with high accuracy by means of a servomotor directly connected to the steering column. The architecture of simulation does not provide a system of motion, because of a fixed base. Three-dimensional images are projected on three 23" screens surrounding the cockpit. Such images immerse the driver in a driving environment realistic that reproduces the road and the surroundings environment (such as houses, vehicles, road signs, trees, etc.) with a high degree of detail. The driving scenarios that can be simulated are varied and include driving at night and in bad weather condition (rain, snow, fog) .The video system ensures a field of front vision of 100°x20° and a total resolution of 5040 x 1050 pixels.

A wrap-around sound system (2 speakers and 1 subwoofer) generates the sounds coming from the driven vehicle (engine, horn, etc.) and from the traffic (other vehicles). A video camera is installed in front of cockpit allowing the researchers to monitor the status of the participants and document the behaviour while driving.

Ultimately, all stimuli provided to the driver can be accurately reproduced so as to ensure the feeling of being in a real vehicle carrying out a real driving task. Furthermore, the quality of video (refresh rate 60 images per second) and audio systems contribute to make the driving experience as realistic and complete as possible.

The DSs use as a driving simulation software SCANNER STUDIO Ver. 1.4 developed by OKTAL, which provides a flexible and effective simulation with an advanced traffic model. The software, in fact, allows to simultaneously handle a traffic of more than 200 different types of autonomous vehicles (motorcycles, cars, trucks, buses, emergency vehicles, bicycles and pedestrians). Each vehicle has assigned driving behaviour rules, which interacts with other vehicles during the simulation. These rules can be replaced by others imposed by the requirements of the scenario (example: control logic). To increase the realism of the driving environment, autonomous vehicles of traffic reproduces the dynamic case by means of the roll (during a planimetric curve)

or pitch (during accelerating or braking phases) and, also, they are able to reproduce the visual signals (marker lights, emergency brake lights, turn lights, fog lights, etc.). In order to generate scenarios with high density of traffic it's possible to create a "swarm" of autonomous vehicles around the interactive vehicle (vehicle controlled by the human driver).

It's, finally, possible to record a huge variety data, such as: data concerned the inputs of the driver's control (steering wheel angle, racing pedals, turn signals, gear, etc.), data related to the vehicle (speed, engine speed, position, etc.), information about autonomous vehicles and, at last data useful to define the behaviour in relation to other vehicles and the road infrastructure. According to the needs of the experiment, it's possible to choose the recording frequency (100Hz max) and the type of data to be recorded.

2.2 Experimental surveys and data collection

In this section the two experimental campaigns have been discussed. In particular the first one involved a large survey (a sample of 100 drivers), while the second one involved 20 drivers. These two experimental campaigns have been fully supported by the DRIVEIN² project.

2.2.1 DriveIn² experiment

The experiment described in this section has been carried out during the DRIVEIN² research project that focuses on defining methodologies, technologies and solutions aimed at capturing driving behaviours (Bifulco, Pariota, Galante, & Fiorentino, 2012). The research project was carried out at the Dipartimento di Ingegneria Civile ed Ambientale of the University of Naples and experimental sessions have lasted from September to October 2012. A sample of 100 participants was selected for experimental purposes, drawn in order to match the population of Italian drivers. The following levels were considered in the stratification of the population:

1. Gender: male and female;
2. Age, 4 classes:

- class 1, from 20 to 24 years old;
 - class 2, from 25 to 40 years old;
 - class 3, from 41 to 64 years old;
 - class 4, over 65 years old;
3. Educational level attained, considered *low* (until high school diploma) or *high* (after graduation).

The admissible combination of the previous features allows the sample to be split over 14 layers (Table 1 below).

<i>Layer</i>	<i>Age</i>	<i>Gender</i>	<i>Educational level</i>	<i>Relative Incidence</i>	<i>Layer Cardinality (over sample of 100)</i>
1	20-24	M	L	0.2*0.429*1	9
2		F		0.2*0.571*1	11
3		M		L	0.3*0.483*0.5
4	H		0.3*0.483*0.5	7	
5	F		L	0.3*0.517*0.5	8
6		H	0.3*0.517*0.5	8	
7		M	L	0.3*0.491*0.5	7
8	H		0.3*0.491*0.5	7	
9	F		L	0.3*0.509*0.5	8
10		H	0.3*0.509*0.5	8	
11		M	L	0.2*0.674*0.5	7
12	H		0.3*0.674*0.5	7	
13	F		L	0.3*0.326*0.5	3
14		H	0.3*0.326*0.5	3	

Table 1 – The sample

The cardinality of each layer depends on the relative incidence on the population given by official data, such as those provided by the ISTAT (*Istituto [Nazionale] di STATistica* – Italian National Statistics Institute), as updated to the latest available year. In order to fill in missing information, we also used data from the DATIS project, carried out by the ISS (*Istituto Superiore della Sanità* – Italian National Health Institute, <http://www.iss.it/chis/?lang=2>) to define the distribution of gender in each single layer.

Having defined the desired cardinality of each layer, we recruited individuals among those responding to an advertisement requesting volunteers

for a study on driving behaviour. Volunteers have been rewarded with 70 euro for participation to both the on-the-road and the driving simulator sessions. Selection was carried out in three steps:

- a) *Contact*, 100 drivers were needed; however, we decided on a preliminary basis to select 150 drivers, distributed according to the desired stratification. Thus a preliminary sample was considered by increasing each of the sample layers by 50%;
- b) *Administration of the pre-selection questionnaire* to those showing interest in participating in the experiment. This questionnaire consisted of four distinct parts: Personal Data; Traffic Locus Of Control Questionnaire (T-LOC); Marlowe Crowne Social Desirability (MCSD); Dangerous Driving (DDDI).

T-LOC, MCSD and DDDI, were administered in random order to respondents, to avoid distortion phenomena, as most people tend to provide the last answers hastily due to the annoyance. Analysis of the questionnaire responses achieved three purposes:

- it confirmed that the respondents belonged to the layer he/she had been contacted for. If the layer was not confirmed, the respondent was switched to the appropriate layer. If, after verification of all respondents a layer was under-represented with respect to the desired classification, the sample was supplemented with new respondents;
 - it classified the respondents by employment status: student (if appropriate for the considered layer, depending on age), employed vs. unemployed, retired (if appropriate for the layer);
 - it divided individuals into clusters. Three categories were created according to responses given to the T-LOC, MCSD and DDDI tests by using two-step cluster analysis (TSC, SPSS Inc. 2001): aggressive drivers; non-aggressive drivers and fatalists; non-aggressive and non-fatalistic drivers;
- c) *Final selection of the sample*: from 150 contacts 100 respondents were selected. For each layer, respondents were selected, obtaining a good balance with respect to employment status and clusters in order to

ensure a representation not lower than 25% for each of three clusters.

The experiment was organised into daily experimental sessions, each consisting of several driving sessions. In order to schedule the driving tests, the drivers made their own reservations for one of the experiment day through a web application. Driving sessions were sequenced every two hours, accommodating driving time and the time required to answer two questionnaires (pre and post-driving). Each daily session involved at most five driving sessions (from 8:30 a.m. to 6:30 p.m.). Since the sun sets before 6 p.m. only in last days of October, similar sunlight conditions were established for each session.

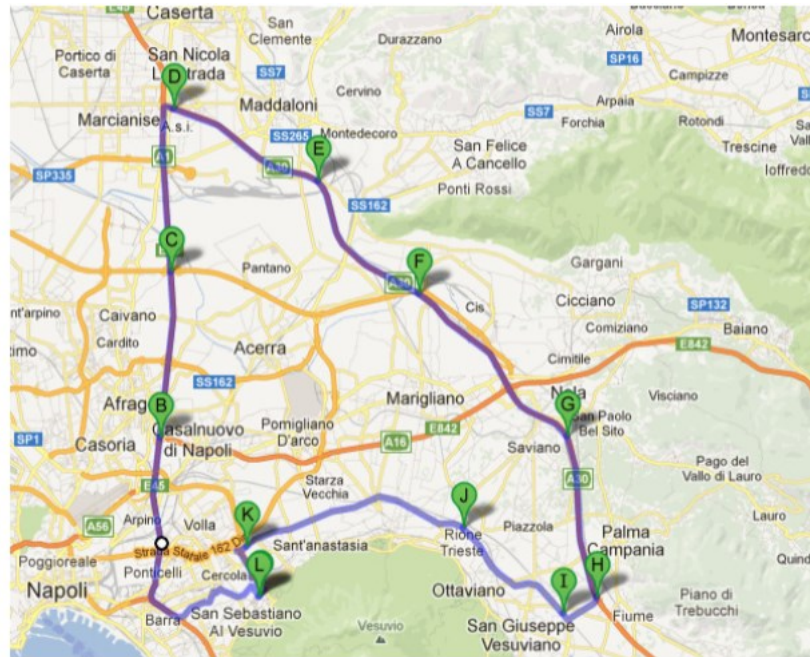


Figure 10 – The experimental path

The following steps were carried out for the observation of each of the 100 drivers:

- each driver drove on the same tour (see Figure 10) by using the instrumented vehicle described previously, detecting the behaviour of the driver with respect to the vehicle ahead and the behaviour of the driver of the vehicle above with respect to the instrumented vehicle;
- the tour is 78 km long, each driving session lasting about one hour; the

route consists in a single loop, evolving over three roads near Naples as in Figure 10 for a total length of 60 km:

- National Highway A1 (from B to D in Figure 10, about 14 km), consisting of a dual carriageway and three lanes for each traffic direction, with a designated speed range of 80-120km/h (speed limit 100 km/h). Here the driver is immersed in a traffic stream that moves at about 100 km/h. Thus natural car-following data are obtained, in the sense that the driver is not asked to perform special tasks;
- National Highway A30 (from D to H in Figure 10, about 30 km), with the same characteristics as National Highway A1, apart from the speed limit (130 km/h). Here the driver interacts with a vehicle taking part to the experiment (corporate vehicle), which carries out several standard manoeuvres. In particular, the driver is asked to perform three approaching manoeuvres with the leader at a constant speed of 80, 100 and 120 km/h;
- The "Vesuvius" State Highway SS 268 (from I to K in Figure 10, about 16 km), consisting of a single carriageway with one lane per traffic direction, at-grade intersections and design speed interval of 60-100 km/h (speed limit 90 km/h). Here the corporate vehicle is not present; however car-following data are obtained.

For part of each of the three main sections a workload experiment was carried out on the drivers, aimed at estimating their mental workload. This refers to the portion of the driver's information processing capacity (or resources) that is actually required to meet requirements in the driving task (Eggemeier, Wilson, Kramer, & Damos, 1991). In particular, Rotated Figures Task (RFT, as in Stanton et al. (1997)) is a self-paced secondary task used to measure performance-based workload. It has been successfully adopted in previous research about workload assessment as a secondary task during driving and its reliability can compensate for its apparent lack of ecological validity. To ensure the effective measure of the spare attentional capacity, this subsidiary task is designed to load the same attentional resources as driving (which are, visual input, spatial processing and manual response; Wickens (2002)). The starting point for driving sessions was Via Gianturco, a major

urban road in the eastern part of Naples, due to the availability of public transport services, thanks to the Gianturco underground station, and quick access to the A1 Highway. Having been met at the beginning of the test-driving route, selected drivers were given a pre-driving questionnaire, comprising the DCQ (Driver Coping Questionnaire) and the DSI-Pre (Driver Stress Inventory-Pre) as described in Matthews et al. (1996) as well as the Italian version of the PANAS (Terracciano, 2003). The tests aimed to investigate the driver's mood prior to driving and to interpret his/her driving behaviour in light of it.

Before point B, a period of acclimatization (from L to B in Figure 10) was introduced, where the driver familiarized him/herself with the instrumented vehicle, to prevent the driving behaviour being biased by lack of familiarity. A final section was introduced (from K to L, see Figure 10) allowing the driver to come back to the starting point. Finally, a post-driving questionnaire was posed to drivers in order to ascertain in what way the driver's mood was influenced by the experiment. The questionnaire comprised the DSI-Post and the NASA-TLX tests adapted for workload (Hart & Staveland, 1988), (Bracco & Chiorri, 2006)). Workload was used, amongst others, to validate an associated experiment carried out using a driving simulator, as presented in Bifulco et al. (2013).

2.2.2 *DriveIn² experiment add-on*

The experiment described in this section is carried out in collaboration with the CNR (*Consiglio Nazionale delle Ricerche – Istituto motori*)

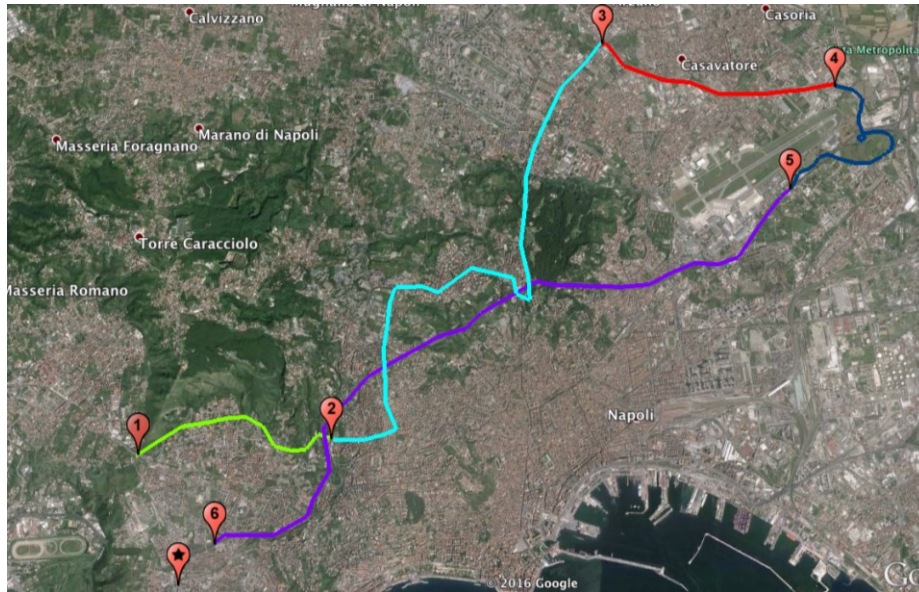


Figure 11 – The experimental path

The route is divided into four main sections, and it is long about 40 km:

- The first section, about 4km long, is made up of the urban freeway “Vomero-Soccavo-Pianura” (Figure 11, green);
- The second section, about 10 km long, is made up of various urban roads that cross several areas of Naples: Vomero, the so called Hospital District of Naples, Colli Aminei, Capodimonte and finally Miano (Figure 11, cyan);
- The third section, 6.5 km long, is the provincial highway 500 “Melito-Scampia” (Figure 11, red);
- The fourth section, about 10 km long, is part of the Autostrada A56, widely known as “Tangenziale di Napoli”, and it is the section in which the higher speeds are reached during the test (Figure 11, purple).

The experiment involved a survey of 20 people, randomly sampled. The sample is composed by 4 females and 16 males, with an average age equal to 29.4 years.

Chapter III – Fuel consumption model

Introduction

In this chapter the fuel consumption model is presented and discussed. The model is specified and calibrated thanks to the dataset obtained from the experimental survey of DRIVEIN² project (paragraph. 2.2.1). The available sample, composed by 100 drivers is randomly split into two groups: the first group is used for specification and calibration purposes, while the second group is employed for testing and validation of the model. Due to the “panel data” typology of the dataset, a linear mixed-effect model is used in order to specify the fuel consumption model. Finally due to the fact that the data are collected using the OBD-II port, as the direct access to the CAN (Controller Area Network) is not allowed, a validation of fuel consumption data collected by the OBD-II port is carried out using a PEMS.

3.1 Linear mixed effect model

For hierarchical data with a single level of grouping, we can formulate the classical LMM (Linear Mixed Model) at a given level of a grouping factor as follows:

$$y_i = X_i\beta + Z_ib_i + \varepsilon_i \quad (12)$$

Where

- y_i is a dependent variable for group i:

$$\mathbf{y}_i \equiv \begin{pmatrix} \mathbf{y}_{i1} \\ \vdots \\ \mathbf{y}_{ij} \\ \vdots \\ \mathbf{y}_{in_i} \end{pmatrix} \quad (13)$$

- X_i is a design matrix for the i-th group:

$$X_i \equiv \begin{pmatrix} x_{i1}^{(1)} & x_{i1}^{(2)} & \cdots & x_{i1}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{in_i}^{(1)} & x_{in_i}^{(2)} & \cdots & x_{in_i}^{(n)} \end{pmatrix} \equiv \begin{pmatrix} x_i^{(1)} & x_i^{(2)} & \dots & x_i^{(p)} \end{pmatrix} \quad (14)$$

- β are the corresponding regression parameters:

$$\beta_i \equiv \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{ij} \\ \vdots \\ \beta_{in_i} \end{pmatrix} \quad (15)$$

- ε_i the vector of the within-group residuals errors:

$$\varepsilon_i \equiv \begin{pmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{ij} \\ \vdots \\ \varepsilon_{in_i} \end{pmatrix} \quad (16)$$

- Z_i and b_i are the matrix of covariates and the corresponding vector of random effects:

$$Z_i \equiv \begin{pmatrix} z_{i1}^{(1)} & z_{i1}^{(2)} & \dots & z_{i1}^{(q)} \\ \vdots & \vdots & \ddots & \vdots \\ z_{in_i}^{(1)} & z_{in_i}^{(2)} & \dots & z_{in_i}^{(q)} \end{pmatrix} \equiv \begin{pmatrix} z_i^{(1)} & z_i^{(2)} & \dots & z_i^{(q)} \end{pmatrix} \quad (17)$$

$$b_i \equiv \begin{pmatrix} b_{i1} \\ \vdots \\ b_{iq} \end{pmatrix} \quad (18)$$

Similar to the design matrix X_i , the matrix Z_i contains known values of q covariates, with corresponding unobservable effects b_i . Moreover:

$$b_i \sim N_q(0, D), \quad \varepsilon_i \sim N_{N_i}(0, R_i), \quad \text{with } b_i \perp \varepsilon_i \quad (19)$$

i.e., the residual errors ε_i for the same group are independent of the random effects b_i . This particular assumption plays the key role in distinguishing a classical LMM from an extended one. In addition, we assume the vectors of random effects and residuals errors for different groups are independent of each other, i.e., b_i is independent of $\varepsilon_{i'}$ for $i \neq i'$.

We also specify that

$$D = \sigma^2 D \quad R_i = \sigma^2 R_i \quad (20)$$

In general, it's assumed that D and R_i are positive-definite, unless stated otherwise.

As we can see, in addition to the fixed-effects parameters β for the covariates used in constructing the design matrix X_i , model (12) includes two random components: the within-group residual errors ε_i and the random effects b_i for the covariates included in the matrix Z_i . The presence of fixed and random effects of known variables justifies the name of model.

In many cases, the random effects included in b_i have corresponding fixed effects, contained in β . Consequently, the matrix Z_i is often created by selecting

a subset of appropriate columns of the matrix X_i . In this case, it is said that the corresponding fixed and random effects are coupled. The notation and theoretical part of this section is presented in the book of Gałecki & Burzykowski (2013).

3.1.1 Specification for all data

In this section, the specification of the single-level LMM, given by (12) - (20), is briefly described for all data.

Let $y \equiv (y'_1, y'_2, \dots, y'_N)$ be the vector containing all $n = \sum_{i=1}^N n_i$ observed values of the dependent variable (observed values of fuel-consumption, in our case). Similarly, let $b \equiv (b'_1, b'_2, \dots, b'_N)$ and $\varepsilon \equiv (\varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_N)$ be the vectors containing all N_q random effects and n residual errors, respectively.

Define matrices

$$\mathbf{X} \equiv \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{N-1} \\ X_N \end{bmatrix} \text{ and } \mathbf{Z} \equiv \begin{bmatrix} Z_1 & 0 & \dots & 0 \\ 0 & Z_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Z_N \end{bmatrix} \quad (21)$$

as we can observe, Z is a diagonal matrix. Moreover, X is of dimension $n \times p$, while Z is of dimension $n \times N_q$. Models (12) - (19), can be written for all data as follows:

$$y = \mathbf{X}\beta + \mathbf{Z}b + \varepsilon \quad (22)$$

With

$$b \sim N_{N_q}(0, \sigma^2 \mathbf{D}) \quad \varepsilon \sim N_n(0, \sigma^2 \mathbf{R}) \quad (23)$$

Where

$$\mathbf{D} \equiv I_N \otimes D = \begin{bmatrix} D & 0 & \dots & 0 \\ 0 & D & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D \end{bmatrix} \quad \mathbf{R} \equiv \begin{bmatrix} R & 0 & \dots & 0 \\ 0 & R & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R \end{bmatrix} \quad (24)$$

Where \otimes denoting the right Kronecker product.

It is worth noting that the particular block diagonal form of matrices \mathbf{Z} , \mathbf{D} and \mathbf{R} , given in (21), (23) and (24), respectively, results from the fact that the single-level LMM, defined by (12) - (20), assumes a particular hierarchy of data of random effects, as explicitly shown in (19). In particular the model assumes that random effects for different groups, defined by level of particular factor, are independent. Informally, we can describe the hierarchy as generated by grouping factors, with one being nested within the other. It's possible, however, to formulate random-effects models by using the representation (22), with non block-diagonal matrices \mathbf{Z} , \mathbf{D} and \mathbf{R} . This is the case, for instance, of models with crossed random effects; but this part will not be described because is out of the scope.

3.1.2 Maximum-Likelihood Estimation

In this section methods to obtain a set o estimates of parameters β , σ^2 , θ_D and θ_R for the classic LMM are presented. For the classical LMM, the marginal mean and variance-covariance matrix of y_i are given as follow:

$$E(y_i) = X_i \beta \quad (25)$$

$$Var(y_i) = \mathbf{V}_i(\sigma^2, \theta, v_i) = \sigma^2 V_i(\theta, v_i) = \sigma^2 [Z_i D(\theta_R) Z_i' + R_i(\theta_R, v_i)] \quad (26)$$

Where $\theta' \equiv (\theta_D', \theta_R')$. To simplify the notation, we will, in general, suppress the use of θ and v_i in the formulae.

From (25) and (26), it follows that, marginally:

$$y_i \sim N_{n_i}(X_i\beta, \sigma^2 Z_i D Z_i' + \sigma^2 R_i) \quad (27)$$

The marginal mean value of the dependent variable vector y_i , is defined by a linear combination of the vectors of covariates included, as columns, in the group-specific design matrix X_i , with parameters β . Moreover, the variance-covariance matrix of y_i , consist of two components. The first one, $\sigma^2 Z_i D Z_i'$, is contributed by the random effects b_i . The second one, $\sigma^2 R_i$, is related to the residual errors ε_i . Hence, strictly speaking, the model employing random effects, specified in (12) - (19), implies a marginal normal distribution, defined in (27) with the variance-covariance matrix of y_i given by (26).

The ML estimation involves constructing the likelihood function based on appropriate probability distribution function for the observed data. The unconditional distribution of b_i and the conditional distribution of y_i given b_i , are not suitable for constructing the likelihood function, because the random effects b_i are not observed. Instead, estimation of LMM is based on the marginal distribution of y_i . In particular the ML estimation is based on the marginal log-likelihood resulting from (27); the log-likelihood can be expressed as follows:

$$\begin{aligned} \ell_{Full}(\beta, \sigma^2, \theta) \equiv & -\frac{N}{2} \log(\sigma^2) \\ & -\frac{1}{2} \sum_{i=1}^N \log[\det(V_i)] - \frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - X_i\beta)' V_i^{-1} (y_i - X_i\beta) \end{aligned} \quad (28)$$

Where V_i , defined in (26), depends on θ and it represents the model-based marginal variance-covariance .

The log-profile-likelihood results from plugging into (28) the estimators of β and σ^2 , given by:

$$\hat{\beta}(\theta) \equiv \left(\sum_{i=1}^N X_i' V_i^{-1} X_i \right)^{-1} \sum_{i=1}^N X_i' V_i^{-1} y_i \quad (29)$$

$$\hat{\sigma}_{ML}^2(\theta) \equiv \sum_{i=1}^N r_i' V_i^{-1} r_i / n \quad (30)$$

Where $r_i \equiv r_i(\theta) = y_i - X_i \hat{\beta}(\theta)$.

By maximizing the log profile-likelihood function over θ , we obtain estimators of these parameters. Plugging $\hat{\theta}$ into (29) and (30) yields to corresponding estimators of β and σ^2 , respectively.

The ML estimates of the variance-covariance are biased. For this reason, the parameters are better estimated using the REML (Restricted Maximum Likelihood) estimation. Toward this end, the log-restricted-likelihood function is considered, in particular:

$$\begin{aligned} \ell_{REML}^*(\sigma^2, \theta) \equiv & \ell_{Full}(\hat{\beta}(\theta), \sigma^2, \theta) + \frac{p}{2} \log(\sigma^2) \\ & - \frac{1}{2} \log \left[\det \left(\sum_{i=1}^N X_i' V_i^{-1} X_i \right) \right] \end{aligned} \quad (31)$$

Where $\hat{\beta}(\theta)$ is specified in (29). From this function, the parameter σ^2 is profiled out by replacing it by the following estimator, corresponding to:

$$\hat{\sigma}_{REML}^2(\theta) \equiv \sum_{i=1}^N r_i' V_i^{-1} r_i / (n - p) \quad (32)$$

This leads to a log-profile-restricted-likelihood function, which only depends on θ :

$$\begin{aligned} \ell^*_{REML}(\theta) \equiv & -\frac{n-p}{2} \log \left(\sum_{i=1}^N r_i' r_i \right) - \frac{1}{2} \sum_{i=1}^N \log[\det(V_i)] \\ & - \frac{1}{2} \log \left[\det \left(\sum_{i=1}^N X_i' V_i^{-1} X_i \right) \right] \end{aligned} \quad (33)$$

Maximization of (33) yields an estimator of θ , which is then plugged into (29) and (32) to provide estimators of β and σ^2 , respectively.

3.2 Validation of OBD fuel consumption

Validation of consumption data is a fundamental pre-requisite of research, as it represents the reference quantity of that research. The OBD port is the unique communication port that has been adopted for obtaining data from the vehicle. From one hand this is a deliberate choice, as we are interested in adopting a variable that can be collected with ease; from the other hand, data obtained from the OBD-II port have lower quality than data that can be directly accessed via the CAN (Controller Area Network) of the vehicle. In our case the direct access to the CAN was not allowed. In case the car-maker allows for the access to CAN (or the solution is directly applied by the car-maker), the quality of the observed data should be a less significant issue.

In our context, the validation of the reliability of collected FC_{inst} has been carried out indirectly. It is straightforward to obtain FC_{inst} from Fuel metering [mg i] by using the following formulation:

$$FC_{inst} \left[\frac{l}{s} \right] = \frac{4 * RPM * Fuel \text{ Metering}}{2 * 1000 * 60 * 825} \quad (34)$$

where:

- 4 is the number of cylinders in the engine;
- RPM is rated for two because we have one injection each 2 RPM;

- Fuel Metering, which represents the milligrams of fuel consumed inside the combustion chamber for each injection cycle measured in that instant;
- 1000 is used to switch from mg to g;
- 60 is the number of seconds in one minute; and
- 825 is the density of diesel fuel (g/l).

Indeed a direct correlation exists between fuel consumption and emissions. Thus, measuring emissions allows for an indirect estimation of fuel consumption, thanks to stoichiometric relationship among fuel burnt, carbon monoxide and carbon dioxide emitted. For this purpose the following formulation are used:

$$FC_{PEMS} \left[\frac{l}{s} \right] = \frac{0.116}{0.8242} \frac{0.273 \times CO_2 + 0.429 \times CO}{100} \quad (35)$$

where CO_2 and CO are respectively the instantaneous mass of carbon dioxide and carbon monoxide expressed in g/s. In other words 0.273 is the carbon mass fraction of CO_2 , CO_2 is the carbon dioxide in g/s, 0.429 is the carbon mass fraction of CO and CO is the carbon monoxide expressed in g/s. Finally for vehicle fuelled with diesel 0.116 is a balance parameter and 0.8242 is the density of tested fuel at 15°C in kg/l. The development of this formulation is an outcome of experiments carried out for the DRIVE IN² project by the Istituto Motori of the National Research Council (CNR) of Italy.

The first step to perform, in order to compare the consumption data come from different instruments, consists on the data tuning. The lowest common denominator of data collected during the transferability experiment is the GPS signal. As the GPS, not only report the exact time, gets the speed profile also; which can be compared with the speed profile collected by the OBD port and with the flow rate of the exhaust gases that has the same form of speed profile for less than a scale factor. Applying this procedure for each driver, involved in this experiment, the data collected by the OBD are in tune with the data collected by the PEMS.

It is here reported absolute percentage error between the fuel consumption collected by the OBD port and the consumption obtained starting from the

emission data collected by the PEMS in the entire driving session. In particular comparison statistics have been reported in Figure 12, and based on the FC_{diff} parameter, defined as:

$$FC_{diff} = \left| \frac{C(T)^{OBD} - C(T)^{PEMS}}{C(T)^{PEMS}} \right| \quad (36)$$

where $C(T)^{OBD}$ is the value of fuel consumption during the entire driving session starting from the OBD data (37); and $C(T)^{PEMS}$ is the value of fuel consumption during the entire driving session estimated indirectly by using emissions measurements, namely:

$$C(T)^{OBD} [l] = \int_0^T FC_{OBD}(t) \cdot dt \quad (37)$$

$$C(T)^{PEMS} [l] = \int_0^T FC_{PEMS}(t) \cdot dt \quad (38)$$

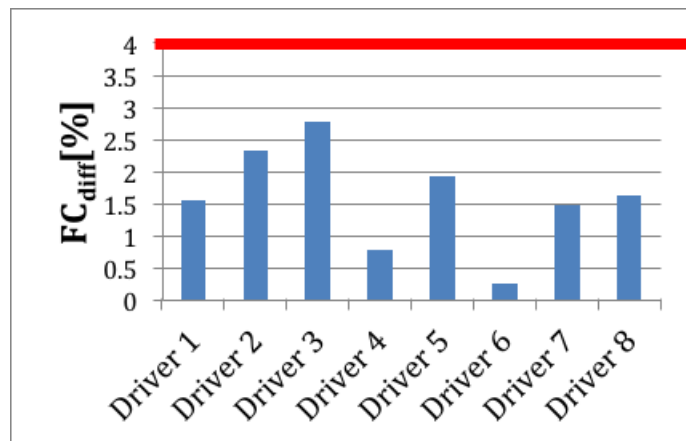
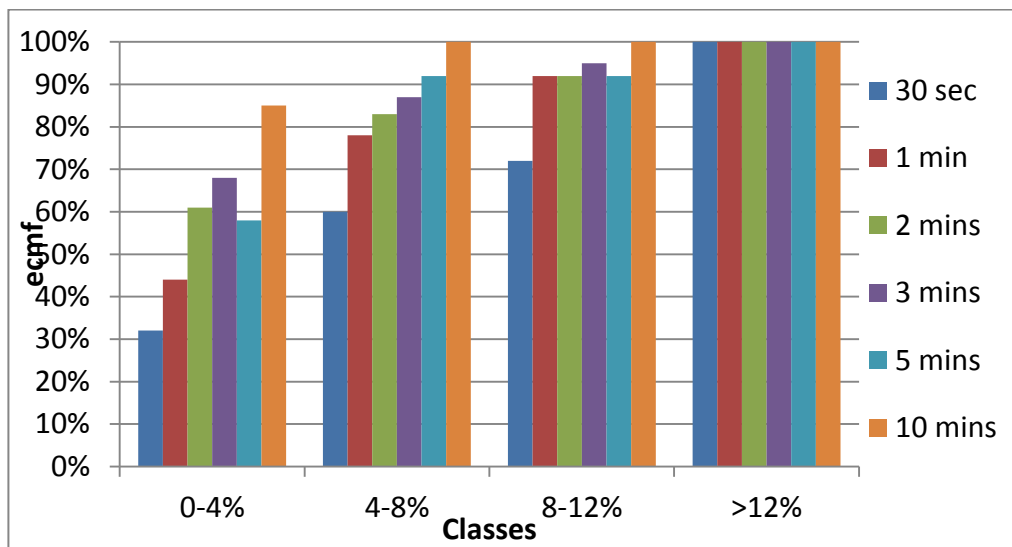


Figure 12 – Percentage differences between aggregate data collected by the OBD port and by the PEMS for each driver.

As shown in Figure 12, the differences between the data collected using the OBD port and the ones estimated by using the PEMS are less than 4%, and confirm the observations by Alessandrini et al. (2012), who carried out a similar experiment (but with different vehicles) in order to evaluate how fuel data collected by the OBD deviate from those collected by the PEMS.

Given that we are not interested in the cumulative fuel consumption of an entire driving session, the data have been aggregated using different time slots (namely 30 sec, 1 min, 2 mins, 3 mins, 5 mins, 10 mins). Different aggregation steps were tested in order to search the best trade-off between a disaggregated enough sampling of data and their accuracy. A fair compromise corresponds to cumulate the data every 3 mins, as reported as example in the figure below (Figure 13).



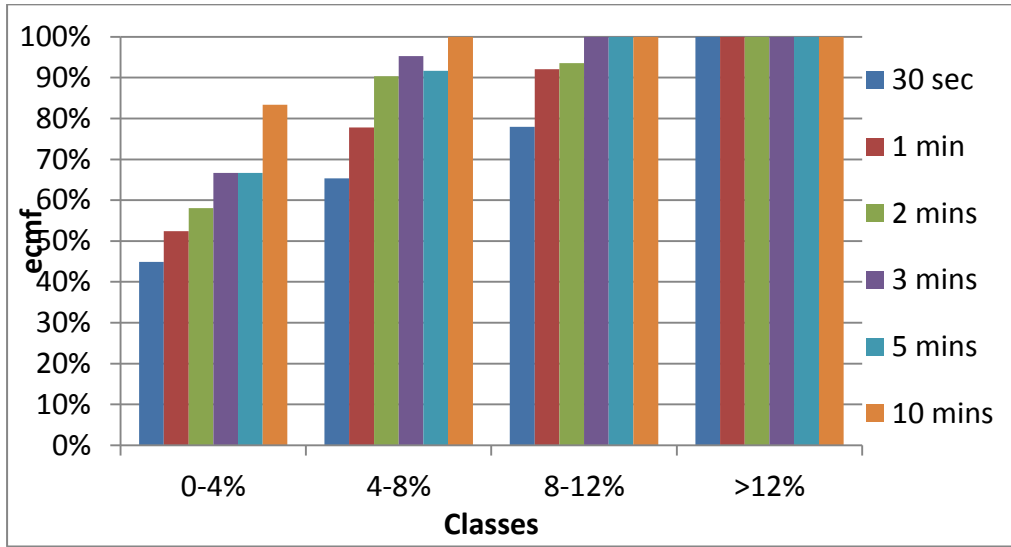


Figure 13 – Empirical cumulative mass functions (ecmf) of gap between data collected by the OBD port and by the PEMS for Driver 6 (on the top) and Driver 3 (on the bottom) using different time slots.

Figure 13 reports the empirical cumulative mass function with different time slots for two drivers, namely driver 6 and driver 3, that are respectively the one exposing the maximum error (FC_{diff} , between OBD data and PEMS data) and the minimum one. It's shown that in more than 70% of the cases the differences are lower than 4% and in 90% of cases the error between the two measurements is lower than 8%.

Of course neither OBD data, nor consumptions estimated from emissions could be considered un-biased. However, the concordance between the two independent instruments gives reasonable confidence on the reliability of measured fuel consumptions, as confirmed by Alessandrini et al. (2012).

Finally, in order to investigate the likely cause of the differences noted at the 3 minutes time-slot, it has been correlated with the value of the median of RPMs calculated in the same 3 minute time slot. Chart in figure 21 shows that low errors (from 0 to 4%) are very frequent at higher RPMs (green bars). Similarly, high levels of error (from 8 to 12%, or even higher) occur at low values of RPMs (red bars). As most of the differences between the two measurements are reported when the value of median RPM is lower than 1000 (idle condition), strictly specking these errors are related at stop phases of

vehicle, that is where (in stop-and-go conditions) our characterization of fuel consumption is not applicable.

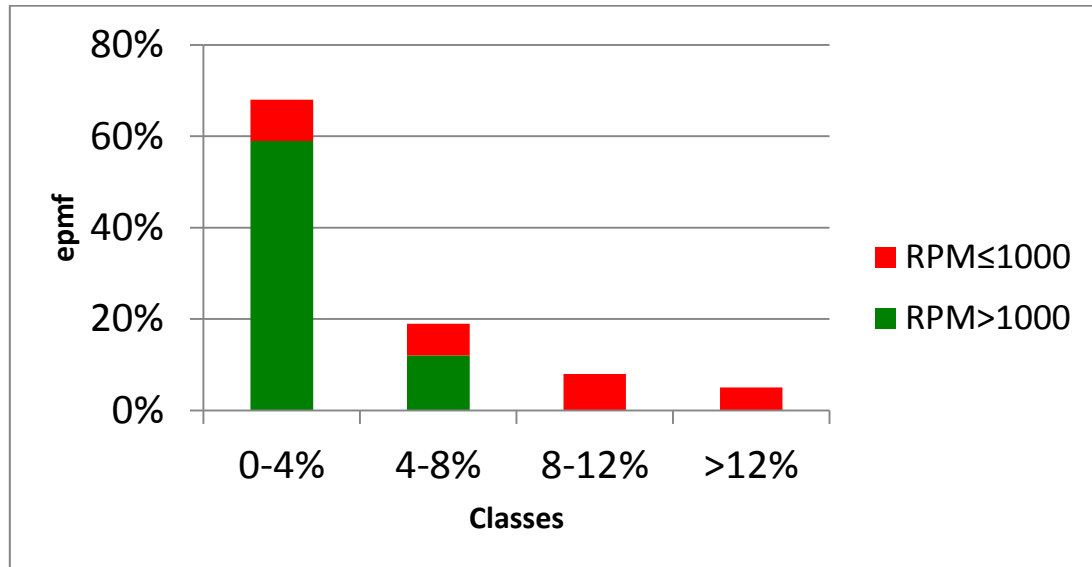


Figure 14 – Empirical probability mass function (epmf) for aggregate data each 3 minutes differentiating for the median value of RPM

3.3 Specification of fuel consumption model

As described in chapter 2 the data carried out during the experimental survey of DRIVEIN² project are used. In particular the sample composed by 100 drivers are randomly divided into two groups: the first one utilized for the characterization of models, the second one for the validation of it.

For what concerns the employed measures, we have used data collected by the on-board CAN (Controller Area Network) of the Instrumented Vehicle, via the OBD (on-board diagnostic) port system, as well as the vector of acceleration along the three axes of motion, provided by an X-Sense inertial measurement unit (IMU). The speed obtained from the OBD has been validated versus the one obtained by a Topcon GPS, adopting the (filtered) GPS speed (sampled at 10 Hz). All data were collected at a frequency of 10 Hz, and were synchronized and recorded on-board. Moreover, to ensure consistent profiles of speed, and acceleration (and also relative speed and spacing with respect to a possible front vehicle), a Kalman filter procedure, described in Bifulco et al.

(2011), was performed on all raw kinematic data used in our modelling. In particular, the analyses refer to the subset of the collected variables mostly related to fuel consumption, focusing especially on the kinematic of the controlled vehicle, and on the driver's interaction with it.

Data taken from the OBD-II port are: speed in km/h, Gas Pedal data (ranging from 0 to 100, represents the travel of gas pedal.), engine revolutions per minute (RPM), fuel metering and the intake of air, which represent respectively the milligrams of fuel consumed and the milligrams of air flowing inside the combustion chamber for each injection cycle measured in that instant. On the other hand, the only information taken from the IMU is the acceleration in m/s^2 along the axis of motion.

To compare all the drivers involved, professional (the FCA test drivers) and non-professional, the analyses refer only to the data collected on the common part of the experimental route.

Though data were recorded at two different frequencies—1 Hz for OBD data and 10 Hz for the rest—the whole data set of measurements was resampled at the lower frequency of 1 Hz, considered as to be fully adequate to the scope of this research. During this sampling procedure, the value of acceleration was approximated with the average of the last ten values.

Fuel consumption is commonly represented in terms of two variables: the instantaneous fuel consumption (FC_{inst}), which expresses the fuel consumption for every second, and the liter per kilometer fuel consumption (FC_{km}), which expresses the fuel consumption in one kilometer if the current motion conditions are maintained stationary.

The FC_{km} can be computed by using the current speed value and instantaneous fuel consumption value (FC_{inst} , eq. (34)) as:

$$FC_{km} \left[\frac{l}{km} \right] = \frac{FC_{inst}}{3600 * Speed} \quad (39)$$

where the speed is expressed in km/h.

In the DRIVEIN² experiment, fuel-metering data collected by OBD were biased by several outliers, and for this reason, some filtering operations were carried out. Filtering operations can be grouped in two phases: a smoothing procedure, and a refinement of the dataset as a whole for the scope of the thesis. In the smoothing phase, a moving average operator was applied to fuel metering; the moving window was fixed at 5 seconds. In the second phase some values were removed from the dataset according to the following criteria:

- fuel metering values lower than 8.62 mgi (these values are evident measurement errors, being below the value of consumption observed in laboratory tests at engine idling speed with Air Conditioner ON);
- instantaneous consumption values higher than 0.12 l/km (these are measurement errors too, being far over the maximum value of fuel consumption furnished by the manufacturer);
- speed values lower than 10 km/h (excluding these cuts off the stop-and-go phase, to which the proposed model does not apply).

The described operations caused the removal of 27% of the rows of our dataset. In this way the final dataset was composed by a full set of variables n-tuple. Despite all the refinement required, the final dataset has about 200 thousand records. Note that the values excluded from the dataset have been discharged on the basis of the value assumed by the observed fuel metering (not the values assumed by the independent variables of the model). Observed fuel metering is the reference value the model should reproduce, and it should be reliable in order to allow estimation of the parameters of the model. The input variables (e.g., RPM, speed, etc.) were not filtered. The only exception to this was for stop-and-go regimes, because these should be modelled independently, and this is not the focus of this thesis. Also note that ADAS that manage the vehicle in the stop-and-go regime (e.g., start and stop automatic systems for engines) are already implemented in many vehicles and that these conditions are not related to the eco-driving attitudes of the drivers.

The model variables were identified by using the stepwise algorithm implemented in the Matlab Statistical Toolbox. This method easily applies to

linear models and can be very effective to be adopted for real-time identification procedures. An obvious correlation of fuel metering with the Gas Pedal and the Intake Air is confirmed by means of Figure 15.

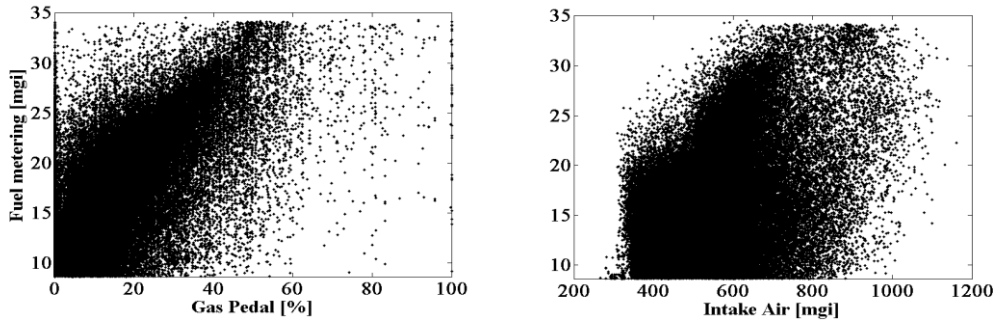


Figure 15 - Dependence of fuel metering on: (lefts side) Gas Pedal; (right side) Intake Air

In addition, considering the law of traction, it is reasonable to presume a dependency of motion from the square of speed (v^2) and acceleration (acc). Indeed these variables were selected as more significant by the stepwise algorithm. In the table below is shown the correlation matrix for the mentioned variables:

	v^2	acc	GasPedal	IntakeAir	FM
v^2	1.0	-0.08	0.37	0.31	0.41
acc	-0.08	1.0	0.27	0.36	0.35
GasPedal	0.37	0.27	1.0	0.42	0.67
IntakeAir	0.31	0.36	0.42	1.0	0.42
FM	0.41	0.35	0.67	0.42	1.0

Table 2 - Correlation matrix

Moreover as for the same circuit different realizations for the same drivers were collected, we are in presence of a linear mixed effect model.

In order to calibrate the model, the R software (R Core Team, 2012) and in particular the *nlme* package (Pinheiro, Bates, DebRoy, & Sarkar, 2015) were used.

The first step to be carried out concerns the calculation of ICC index (Intra Class Correlation index), it represents a measure of reliability, or dependence among individuals (Kreft & DeLeeuw, 1988). It allows us to assess whether or not the random effect is present in the data. To get the ICC, first a null model was created; which would include just the intercepts (fixed and random):

$$FM_{Null} \sim 1 + (1|Id) \quad (40)$$

Obtaining the results reported in the following figure (Figure 16)

```
Linear mixed model fit by REML ['lmerMod']
Formula: FM ~ 1 + (1 | Id)
Data: db100

REML criterion at convergence: 538227.2

Scaled residuals:
    Min       1Q   Median       3Q      Max
-1.8373 -0.8176 -0.0974  0.6543  3.2682

Random effects:
 Groups   Name                Variance Std.Dev.
 Id      (Intercept)    0.7464    0.864
 Residual                    30.8555    5.555
Number of obs: 85850, groups: Id, 50

Fixed effects:
              Estimate Std. Error t value
(Intercept)  17.2669     0.1236   139.7
```

Figure 16- Screenshot of the results for the null model (40)

Next, adding the random effect variance estimates and then dividing the random effect of subject's variance estimate by the total variance estimate, it's possible to estimate the ICC coefficients:

$$ICC = \frac{30.855}{0.7464 + 30.855} = 0.98$$

So, the ICC is 0.98, denoting the presence of random effects.

In order to introduce random effects, first of all it's presented below a box-whisker plot of FM for each subject.

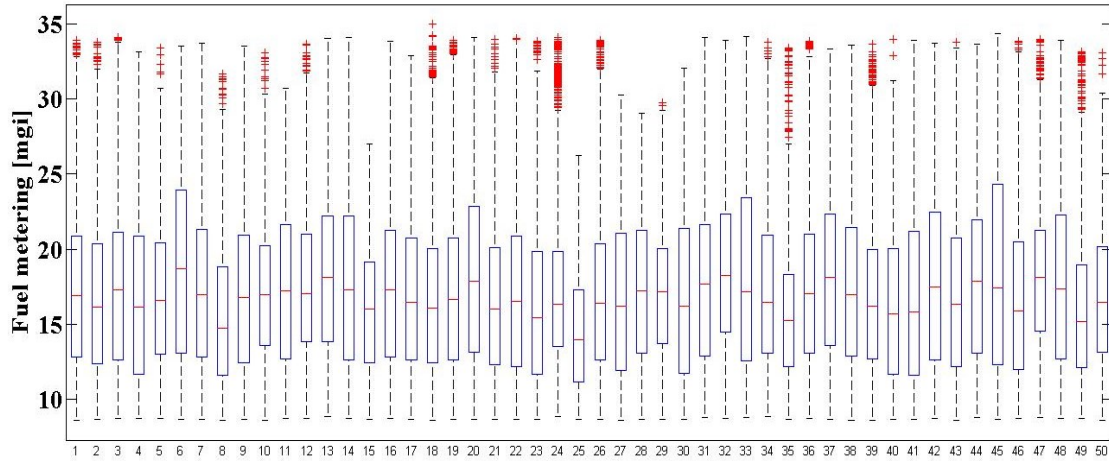


Figure 17 – Box whisker plot of Fuel metering vs subjects

As shown in the Figure 17, it's necessary introduce a random intercepts for subjects (Id); than for what concerns random slopes for subjects, GasPedal and acceleration were introduced because of during the test the subjects had to drive following defined speed tasks.

Finally, as fixed effects the acceleration, the power of speed, the GasPedal and the IntakeAir variables were entered into the model. As random effects, we had intercept for subjects (Id), as well as by-subject random slopes for the effects of GasPedal and acceleration.

$$FM_{mgi} \sim 1 + v^2 + acc + GasPedal + IntakeAir + (1 + acc + GasPedal|Id) \quad (41)$$

In the Table 3 the results are reported:

	Parameter	FM _{mgi}
Model Label		(41)
AIC		459741.8
BIC		459854.1

LogLik		-229858.9
Fixed effect		
Intercept	β_0	6.98 (0.19)
v^2	β_1	0.78 (0.01)
acc	β_2	13.01 (1.34)
GasPedal	β_3	0.23 (0.01)
IntakeAir	β_4	-0.000468 (0.0001)
reStruct(Id)		
SD(b _{i0})	$\sqrt{d_{11}}$	1.23 (0.98;1.53)
SD(b _{i1})	$\sqrt{d_{22}}$	9.48 (7.80;11.53)
SD(b _{i2})	$\sqrt{d_{33}}$	0.07 (0.06; 0.08)
Scale	σ	3.505 (3.488; 3.521)

Table 3- Parameters of lme model (41)

In particular the following parameters are specified:

- AIC (Akaike's Information Criterion) is a statistical model fit measure. The index takes into account both the statistical goodness of fit and the number of parameters that have to be estimated to achieve this particular degree of fit, by imposing a penalty for increasing the number of parameters. Lower values of the index indicate the preferred model, that is, the one with the fewest parameters that still provides an adequate fit to the data,

$$AIC = 2k - 2\ln(L)$$

- BIC (Bayesian Information Criterion) is an information criterion for model selection among a finite set of models; the model with the lowest BIC is preferred. When fitting models, it is possible to increase the likelihood by adding parameters, but doing so may result in overfitting. Both BIC and AIC resolve this problem by introducing a penalty term for the number of parameters in the model; the penalty term is larger in BIC than in AIC,

$$BIC = k \ln(n) - 2\ln(L)$$

Where k is the number of regressors, L represents the maximized value of the likelihood function of the estimated model, and finally n is the number of observations;

- LogLikelihood represents the likelihood function, or even its logarithm.

For what concerns fixed effects, the value of beta vectors, with the relative standard deviation in parenthesis, are reported.

Whereas related to random effects, supposing a normal distribution with zero mean and diagonal variance-covariance matrix; thus random intercepts and random slopes have different variances and are uncorrelated; herein the standard deviation values for each ones are reported ($\sqrt{d_{11}}$, $\sqrt{d_{22}}$, $\sqrt{d_{33}}$) with the relative 95% CI. Finally, under the hypothesis the residuals follow a normal distribution with zero mean and diagonal variance-covariance matrix, too; it's reported the standard deviation (σ) and the 95% CI.

Indeed, as for dataset with high sample size, the p-value related to variable isn't meaningful; the standardized beta values for fixed effects of model (41) are reported (see Table 4). The standardized coefficients are the estimates resulting from a regression analyses that have been standardized so that the variance of dependent and independent variables are equal to 1. Thus the coefficients are mutually comparable, as the variables are measured in different measurements units, the standardized of the coefficient is usually done to answer the question of which of independent variables have a greater effects on the dependent one in a multiple regression analyses:

	<i>Standardized beta</i>
v^2	0.27
acc	0.28
GasPedal	0.49
IntakeAir	0.005

Table 4 – Standardized beta for model (41)

As shown in the Table 4, among the identified key parameters, Gas Pedal is the more important variable, while Intake Air is the least important one. As stated before, the aim of this work is to develop a model that can be used in an integrated simulation environment. The need for adopting independent variables that are very easily estimated by simulation platforms or collected from on-board low-cost devices has been clearly stated in Chapter 1. Typically,

Intake Air is difficult to estimate in standard models of vehicle dynamics with enough accuracy. Thus we proceeded to the specification of another model, FM_{reduced} . The reduced model carries the same variables mentioned before except for Intake Air, namely:

$$FM_{\text{mgi}} \text{reduced} \sim 1 + v^2 + \text{acc} + \text{GasPedal} + (1 + \text{acc} + \text{GasPedal} | Id) \quad (42)$$

The same calibration procedure was repeated and also for this case calibration results have been reported in Table 3.

	Parameter	FM_{mgi}
<i>Model Label</i>		(42)
AIC		459751.2
BIC		459826.1
logLik		-229867.6
<i>Fixed effect</i>		
Intercept	β_0	6.80 (0.18)
v^2	β_1	0.77 (0.01)
acc	β_2	12.95 (1.34)
GasPedal	β_3	0.23 (0.01)
<i>reStruct(Id)</i>		
SD(b_{i0})	$\sqrt{d_{11}}$	1.21 (1.07; 1.36)
SD(b_{i1})	$\sqrt{d_{22}}$	9.41 (7.69; 11.49)
SD(b_{i2})	$\sqrt{d_{33}}$	0.07 (0.05; 0.08)
<i>Scale</i>	σ	3.505 (3.49; 3.52)

Table 5 – Parameters of lme model (42)

And for what concerns standardized coefficients:

	<i>Standardized beta</i>
v^2	0.27
acc	0.28
GasPedal	0.49

Table 6 – Standardized beta for model (42)

It is worth noting that, in accordance with the previous consideration about

the standardized coefficients in Table 1, excluding Intake Air represents an acceptable approximation. Indeed further analyses were performed; in particular a logLikelihood Ratio test between the two models specified was performed in order to demonstrate that the use of the simplified model is fully justified. This is possible as the two models have the same random effects and they differ only for the fixed effects; indeed the ML estimation method is meaningful than REML one to perform the logLikelihood ratio test.

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
FM	1	9	459728.1	459812.4	-229855.0			
FM _{reduced}	2	8	459736.4	459811.3	-229860.2	1 vs 2	10.29113	0.0013

Table 7 - LogLikelihood Ratio test for lme models (41) - (42)

In statistics, a likelihood ratio test is a statistical test used to compare the goodness of fit of two models, one of which (the *null model*) is a special case of the other (the *alternative model*). The test is based on the likelihood ratio, which expresses how many times more likely the data are under one model than the other. This likelihood ratio, or equivalently its logarithm, can then be used to compute a *p-value*, or compared to a critical value to decide whether to reject the null model in favour of the alternative model.

It is inferred that the IntakeAir variable is a redundant variable, moreover it penalizes the full model since the BIC criterion information is lower for the reduced model.

The variance-covariance matrix of random effects, D defined in (20) is a diagonal matrix equal to:

	Intercept	acc	GasPedal
Intercept	1.46	0	0
acc	0	88.50	0
GasPedal	0	0	0.004

Table 8 – Variance-Covariance matrix of random effects for lme model (42)

Whereas the variance-covariance matrix of random residuals for each subject, R_i defined in (20) is:

$$\hat{R}_i = \hat{\sigma}^2 I_N = 12,285 \times I_N$$

Indeed, the best of the functional forms within those introduced by Lee et al. (2011), corresponds to the following model:

$$FM_{mgi} reduced1 \sim 1 + GasPedal + (1 + acc + GasPedal|Id) \quad (43)$$

which seems to be less convincing than reduced (42); as shown in the logLikelihood Ratio test reported below:

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
FM _{reduced1}	1	6	471159.9	471216.0	-235573.9			
FM _{reduced}	2	8	459736.4	459811.3	-229860.2	1 vs 2	10.29113	<0.0001

Table 9 - LogLikelihood Ratio test for lme models (43) - (42)

It is worth noting that in the original work the Throttle was used instead of the Gas Pedal as independent variable, but this difference seems to be not enough to justify the great decrease in the performance of the model.

Finally, the last step concerns the check of following hypothesis:

1. Normal distribution of independent variables,
2. Assumption on residuals:
 - a. Zero mean: $E(\epsilon_i) = 0$,
 - b. Homoschedasticity: $VAR(\epsilon_i) = \sigma$
 - c. Normality of residuals errors.

Starting from the first assumption, as the dataset has a high size, the Central Limit Theorem implies that, as the sample size grows, the independent variable distribution approaches normality.

Indeed, for what concerns the assumptions on the residuals, herein the residuals' histogram and the plot of estimated variable vs residuals are reported. Moreover, due to the structure of LMM, various type of raw residuals can be defined. In particular, it's important remark the different between marginal residuals and conditional residuals.

The conditional residuals follow from the conditional mean representation, and are defined as:

$$\hat{\varepsilon}_{(c)i} \equiv y_i - X_i\hat{\beta} - Z_i\hat{b}_i \quad (44)$$

Whereas the marginal residuals result from the marginal mean representation, as:

$$\hat{\varepsilon}_{(m)i} \equiv y_i - X_i\hat{\beta} \quad (45)$$

Santos Nobre and da Motta Singer (2007) argue that marginal residuals are pure, in the sense that the residuals are a function of only the marginal errors, $\varepsilon_{(m)i} \equiv y_i - X_i\beta$, which they are supposed to estimate. On the other hand, the conditional residuals, which estimate the residual errors ε_i , are confounded with the random effects b_i , because the residuals are a function of ε_i and b_i . For this reason, they recommend to use the plots of the conditional residuals against the estimated conditional means to detect heteroscedasticity of residual errors and the marginal residual in order to check the normality of the residual errors (Figure 18).

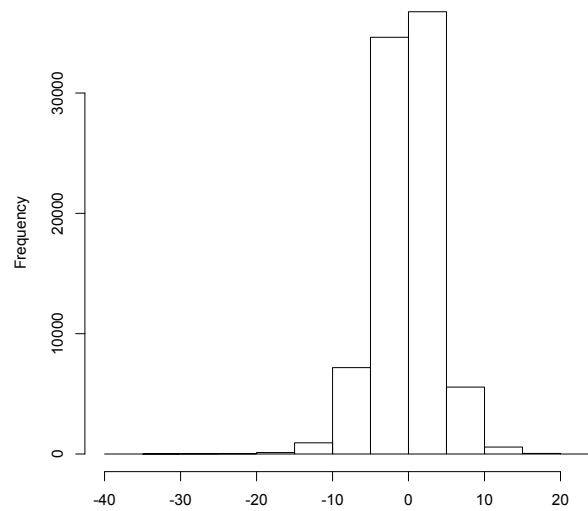


Figure 18 – Histogram of marginal residuals

Finally the Figure 19 shows the results about the homoscedasticity hypothesis

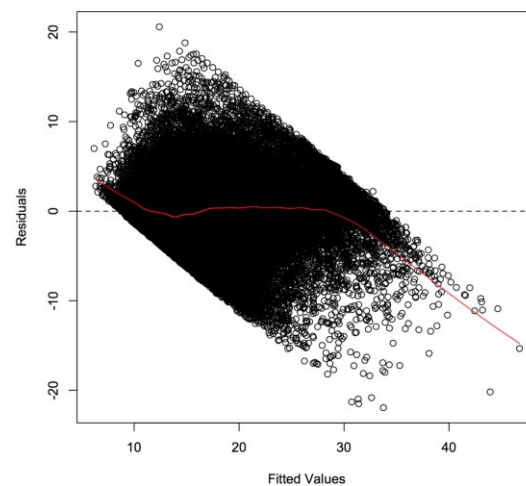


Figure 19 – qq plot conditional residuals vs fitted values

As the residuals are heteroscedastic, a model with random intercepts and random slopes and with heteroscedastic residual errors is introduced, in particular:

$$FM_{mgi}reduced2 \sim 1 + v^2 + acc + GasPedal + (1 + acc + GasPedal | Id) \quad (46)$$

With:

$$R_i = \sigma^2 \mathbf{x} \begin{bmatrix} (Fitted_{i1})^{2\delta} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (Fitted_{in})^{2\delta} \end{bmatrix}$$

in particular the variances are a power function of the mean value:

$$\sigma_{it} = \sigma \lambda_{it} = \sigma \lambda(\mu_{it}) \quad (47)$$

where $\mu_{it} \equiv \beta_0 + \beta_1 v^2 + \beta_2 acc + \beta_3 GasPedal$ is the predicted (mean) value of FM_{it} .

	Parameter	FM _{mgi}
<i>Model Label</i>		(46)
AIC		456166.8
BIC		456251.2
logLik		-228074.4
<i>Fixed effect</i>		
Intercept	β_0	6.98 (0.17)
v^2	β_1	0.72 (0.01)
acc	β_2	13.07 (1.33)
GasPedal	β_3	0.24 (0.01)
<i>reStruct(Id)</i>		
SD(b _{i0})	$\sqrt{d_{11}}$	1.17 (1.03; 1.32)
SD(b _{i1})	$\sqrt{d_{22}}$	9.36 (7.30; 11.42)
SD(b _{i2})	$\sqrt{d_{33}}$	0.07 (0.06; 0.08)
<i>Variance Function</i>		
Power(Fitted ^δ)	δ	0.63 (0.61; 0.65)
<i>Scale</i>	σ	0.58 (0.55; 0.61)

Table 10 – Parameters of lme model (46)

The variance-covariance matrix of random effects, D, is a diagonal matrix equal to:

	Intercept	Acc	GasPedal
Intercept	1.38	0	0
acc	0	86.62	0
GasPedal	0	0	0.005

Table 11 - Variance-Covariance matrix of random effects for lme model (46)

And the variance-covariance matrix of random residuals, for each subject, R_i , is a diagonal matrix as shown below:

$$\begin{aligned}\hat{R}_i &= \hat{\sigma}^2 \times \begin{bmatrix} (Fitted_{i1})^{2\delta} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (Fitted_{in})^{2\delta} \end{bmatrix} \\ &= 0,33 \times \begin{bmatrix} (Fitted_{i1})^{2(0.63)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (Fitted_{in})^{2(0.63)} \end{bmatrix}\end{aligned}$$

Verified the model assumption, the next step consists on model validation.

In particular some statistical considerations have also been carried out with reference to the performance of the model. In Figure 20, the empirical probability mass function (epmf) and the empirical cumulative mass function (ecmf) of the absolute percentage error, ERR_{inst} , have been depicted:

$$ERR_{inst}(t) = \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (48)$$

ERR_{inst} represents the percentage deviation between the predicted value by the model, \hat{x}_i , and that observed, x_i , at the instant i of FM_{mgi} .

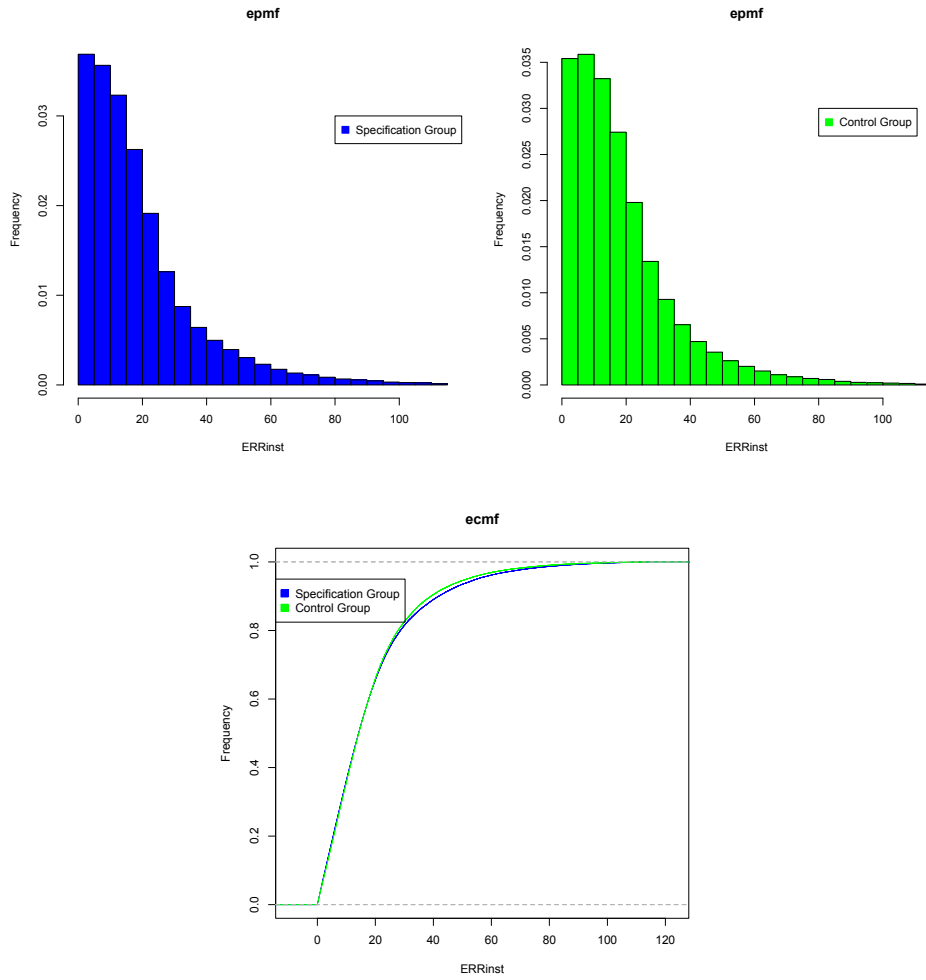


Figure 20 – Empirical probability mass function (epmf) and empirical cumulative mass function (ecmf) vs the percentage error (ERR_{inst}) of the two sub-samples.

Differences in the model outputs and observed values are lower than 30% in more than 80% of the cases (Figure 20).

Moreover, the following statistical indicators have been reported in Table 12:

- RMSE (Root Mean Square Error), which is the ratio between the inner deviance and the total numerosity, n :

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n}} \quad (49)$$

- MAPE (Mean Absolute Percentage Error), which is a measure of accuracy of a method for constructing fitted time series value:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (50)$$

	Specification Group	Control Group
RMSE	3.927	3.778
MAPE	0.194	0.188

Table 12 – RMSE and MAPE computed for the two sub-samples

Another interesting point of comparison of model performances concerns the performance in terms of aggregate fuel consumption $C(T)$. Aggregated data is intended here as the temporal integration of FC_{inst} , defined in equation (34), in the total length of the driving session T , as detailed in the previous equation (37) Even in this case a discrepancy parameter is defined:

$$ERR_{aggr}(t) = \left| \frac{C(T) - \hat{C}(T)}{C(T)} \right| \quad (51)$$

where $\hat{C}(T)$ and $C(T)$ are the aggregate fuel consumption using the values predicted by the model and the observed ones, respectively. Figure 21 shows the empirical probability mass function (epmf) and empirical cumulative mass function (ecmf) of the values assumed by ERR_{aggr} variable.

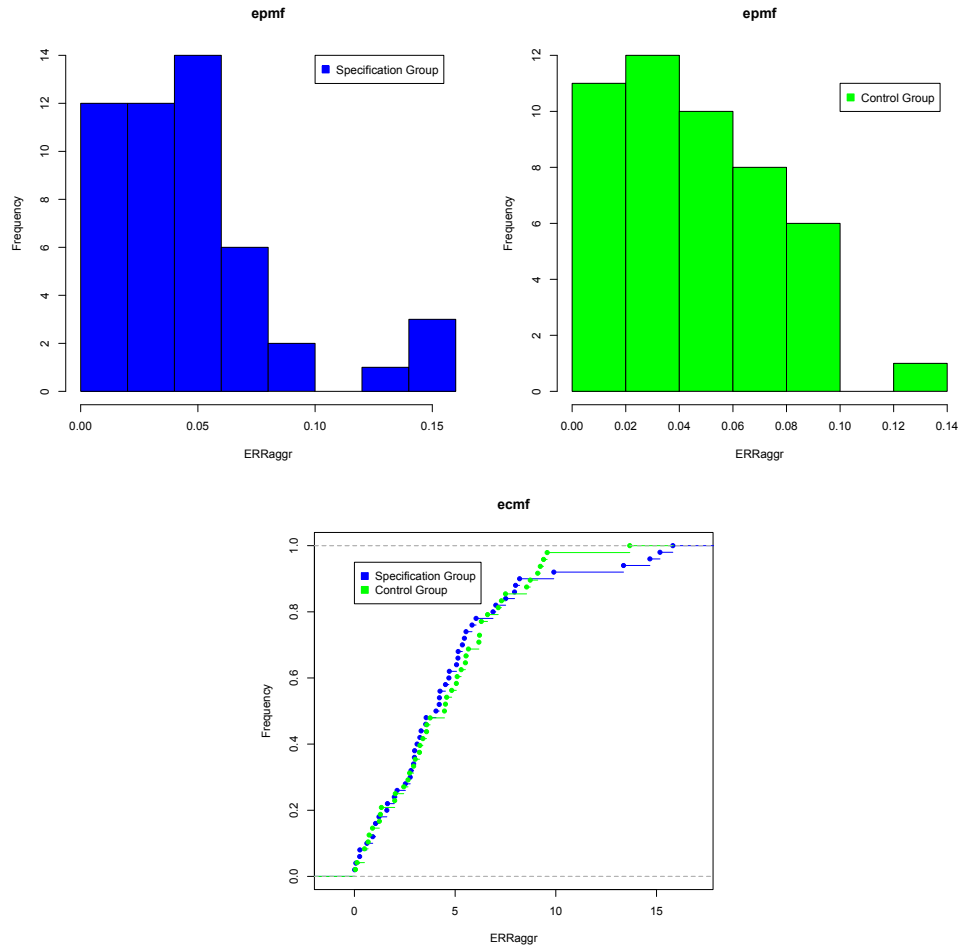


Figure 21 – Empirical probability mass function (epmf) and empirical cumulative mass function (ecmf) of the aggregate error (ERR_{aggr}) for the calibration and the control sub-samples.

It is found that in aggregate terms differences between the model outputs and observed values are lower than 8% in more than 85% of the cases (Figure 21).

The RMSE and MAPE indicators are reported below (Table 13):

	Specification Group	Control Group
RMSE	0.177	0.154
MAPE	0.047	0.045

Table 13 – RMSE and MAPE for aggregate measurements

Chapter IV – Implementation of ADAS in MIL and DIL environment

Introduction

MIL and DIL simulations are carried out in order to validate the technical applicability of the solution in real-time conditions and to test the acceptability of the solution for the drivers.

In order to perform this task, an *eco*-ACC device has been implemented. To carry out a more realistic simulation a vehicle dynamics model was coupled with the ADAS model. The purpose of the implemented vehicle dynamics model is to simulate, in real time, the vehicle's dynamic behaviour, thanks to continuous integration of the balance equations regulating the longitudinal and lateral vehicle motion.

In order to make the simulations as close to reality as possible, the implemented mathematical model also includes working dynamics for wheels, tyres, the engine and braking system, as well as a control system for automatic gearbox activation. Finally, in order to optimize consumption performance during driving task, the specified and validated fuel consumption model (paragraph 3.3) is embedded into the system discussed in the following.

4.1 The implemented model in MIL

The *eco*-drive system implemented is in particular an *eco*-ACC (*eco*-Adaptive Cruise Control), as claimed in the introduction section. Rakha (2015) defined it as a combination of car-following models with ECC (*eco*-Cruise Control) systems; the aim of an ECC system is to find an optimal vehicle velocity profile using an optimal control theory to minimize the vehicle fuel consumption level. It's worth observing that in our case the ADAS system implemented is an *eco*-ACC with human-like characteristics; able to update the

library parameters in order to adopt a driving style not only human-like (Bifulco, Pariota, Simonelli, & Di Pace, 2011) but also able to reduce fuel consumption taking into account the actual stimuli the drivers receive from the traffic. Some details about the operating principle of the adopted solution is shown in the blocks diagram below (Figure 22).

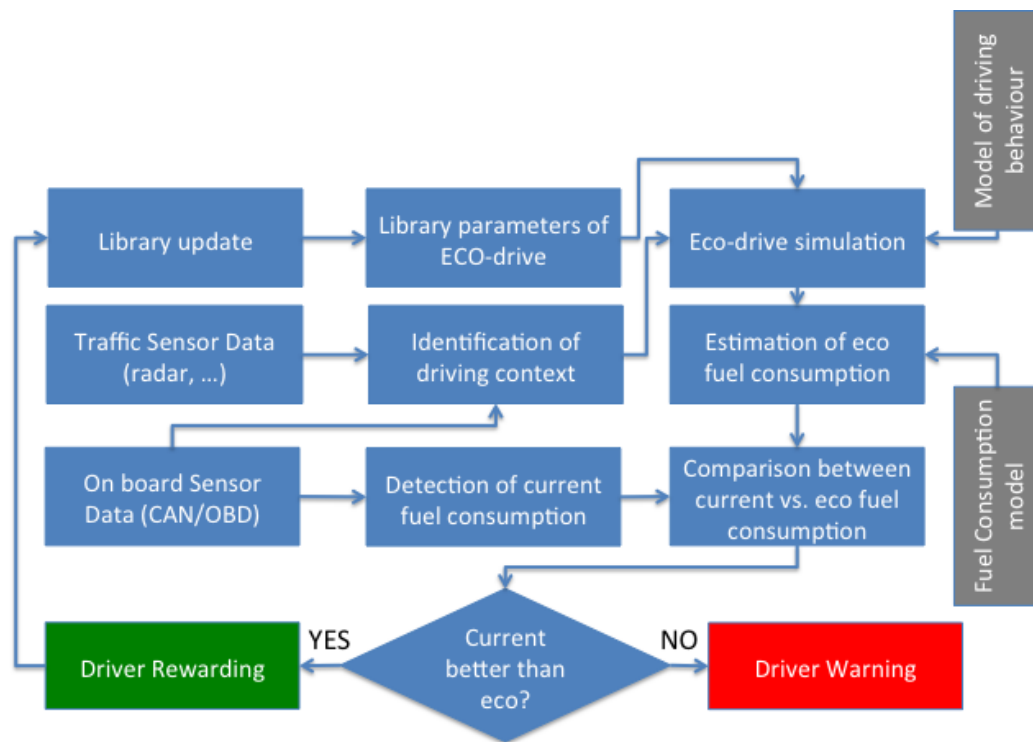


Figure 22 – Scheme of eco-drive system

As shown it's a feedback control system, in fact the implemented system is able to update the library parameters in order to adopt a driving style not only human-like but also oriented to reduce fuel consumption taking into account the actual traffic conditions. In particular it is a system able to observe the driving behaviour and to identify in real-time the parameters of corresponding behavioural model; at the same time the current fuel consumption is observed. Moreover, it is simulated the driving behaviour that would have been applied by an "eco" driver under the same stimuli received from the traffic, as well as the corresponding fuel consumption, as estimated from the consumption model. The "eco" driver is characterised by the behavioural parameters extracted from

the library of the ones estimated in real-time for drivers that have shown actual low-levels of fuel-consumptions. If the actual driver shows observed fuel-consumption that is lower than the eco-estimated one, the behavioural parameters of the current driver are stored in the “eco” library and the driver is gratified; otherwise the driver is warned in order to adopt a more “eco” driving behaviour.

The proposed solution is based on the sub-models listed below:

- the specified and validated fuel consumption model able to estimate the consumption starting from kinematic and driving parameters detected in real- time (paragraph 3.3),
- a reference model that identifies the “target” driving styles taking into account the actual boundary conditions (traffic conditioning), and finally
- a model for the evaluation of the discrepancy between driving style applied and ideal driving style.

In the Figure 23 below the Simulink blocks that implements the solution in the MIL environment are presented and discussed.

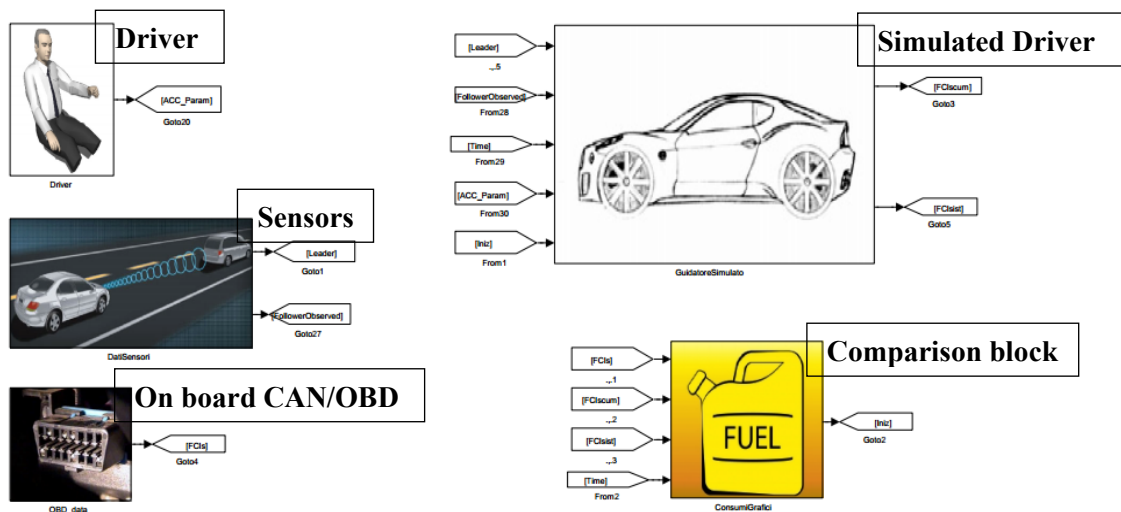


Figure 23 – Architcture implemented in Matlab/Simulink

The implemented system is composed by the following blocks:

- Driver, it consists in two main parts, the first one related to the parameters library of drivers that have reproduced an *eco-driving* style, the second one related to the *learning-phase* introduced and explained in Bifulco et al. (2011). The latter, in fact, is able to calibrate in real-time the ACC model parameters. The ACC model proposed by the authors is a linear stimuli-response model; and it is based on a simple time-discrete linear model, able to relate the instantaneous speeds of leader and follower and their spacing with the target variation of spacing that the followers desired for the next time-step. This model will be briefly explained in the “Simulated Driver” block, in particular, for what concerns the *learning-phase* the estimation process for the behavioural model consists in finding the regression parameters that minimize the distance between the observed vehicle kinematics and the kinematics reproduced by the model. In other words, the calibration algorithm has been implemented; the inputs provided by the “Sensor” block are the instantaneous observed kinematics conditions of leader and following vehicles (spacing, velocity and acceleration), and finally the instantaneous distances between the two vehicles. In order to store and update, or not, the library parameters with those observed, the ratio between the observed fuel consumption and simulated ones is provided by the “Comparison” block. Then the output signals provided by this block are the regression parameters that feed the ACC module of “Simulated Driver” block. This module differs from those presented by authors in the time slot adopted. In particular the system has a memory of only one minute of preceding time history; strictly speaking the dataset utilized for the calibration of regression parameters is composed by only 600 records (refers to a sample frequency of 10 Hz).

- Sensors, it is a block able to get information about the traffic context, in particular it is a block that gives, as outputs, information about the kinematic motion of a vehicle leader and follower ones. As previously described, these outputs represent the inputs of “Driver” block; but also the input of “Simulated Driver” block due to the fact that each one-minute the system refreshes the observed initial conditions. The acquisition frequency is equal to 10 Hz.

- OBD port, in order to collect the instantaneous fuel consumptions observed in real environment with a samples frequency of 10 Hz. These outputs become the inputs of “Comparison” block.

- Comparison block, in this block it's presented the logic that aggregates the data each 1 minute and compare the two values obtained (actual consumption vs. estimated consumption by the *eco-driver*). This part has been developed using the *Stateflow* Toolbox of Matlab/Simulink software. This block presents as inputs the fuel consumptions observed and ones calculated by the system; while as output provides da ratio between the observed and calculated fuel consumption, in order to update, or not, the library parameters of the *eco-ACC* presented in the “Driver” block. Major details about this part will be explained in the next section (paragraph 4.2).

- Simulated Driver, it's the core of the system developed, it's composed by different sub-systems explained in detail and shown below. This sub-system is composed by the ACC human-like logic, the vehicle dynamic model and finally the fuel consumption model illustrated in chapter 3.

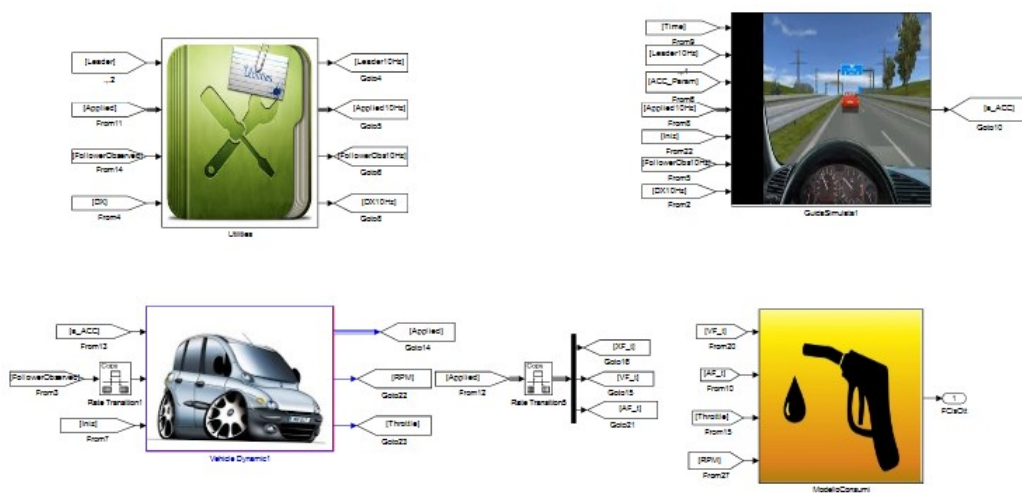


Figure 24 – ‘Simulated Driver’ sub-system

The inputs of this system are: the observed kinematics motions of follower and leader derived from “Sensor” block, the ACC parameters derived from “Driver” block; and it provides the instantaneous estimated fuel consumptions by *eco-driver* as output.

The “simulated driver” is based on the human like ACC developed and validated by Pariota during his PhD (2013). The system is based on a simple time-discrete linear model, able to relate the instantaneous speeds of leader and follower and their spacing with the target variation of spacing that the followers desired for the next time-step. In particular this block refers to the so called *running-mode*; it is composed by the *sampler*, *profiler* and *tutor* blocks defined by the author. A briefly description is herein given: the *sampler* is the main responsible of ACC control logic, the *profiler* translates the driving behaviour produced by the *sampler* into admissible (continuous) kinematic profiles, suitable to be applied to the controlled vehicle; and finally, the *tutor* ensures that the reference driving trajectory obtained by the *profiler* is ultimately safe. The major effort carried out consists in the portability of this system from an older release of Matlab/Simulink to a newest one. The block presents as inputs the kinematic characteristics of the leader vehicle and the kinematic characteristic calculated by the system and the ACC parameters of the system. Furthermore, the model provides as output a desired acceleration that becomes the input of the vehicle dynamics block, that is in charge to evaluate if the required acceleration can be really applied by the vehicle. In Pariota (2013) this task was demanded to a dedicated software of vehicle dynamics model. Indeed, the model of the “vehicle dynamics” deals with a more general vehicle-behaviour model. Besides the ‘pure’ vehicle dynamics equations, also the modelling of some other essential vehicle components (such as the engine, gearbox, braking system and others) was performed. In order to make the vehicle dynamics model clearer and more readable, it was organized into different sub-systems (Russo Spena, Timpone, & Farroni, 2015) as shown in Figure 25.

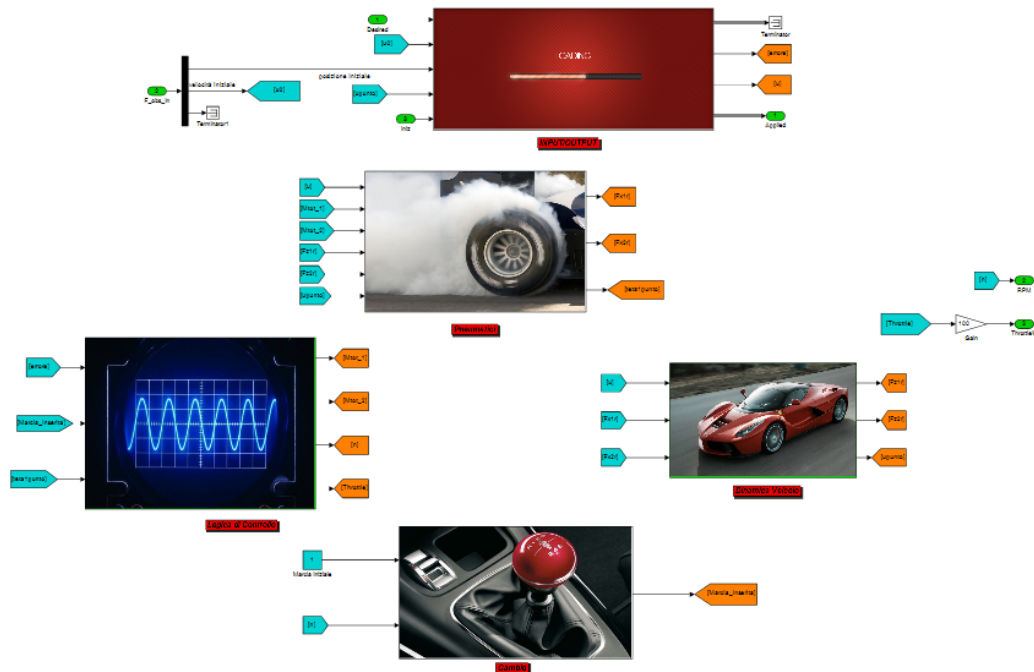


Figure 25 – Screenshot showing the vehicle dynamics Simulink model

The main subsystem contains all the equations describing the vehicle dynamics: the Longitudinal Balance Equation, Lateral and Yaw Vehicle Dynamics and Vertical Load Determination, taking into account load transfers by means of the handling diagram/phase plane approach described in Farroni et al. (2013) and of the influence of roll stiffness (Farroni F. , Russo, Russo, Terzo, & Timpone, 2012).

In the input/output subsystem all the model's exchange parameters with the other subsystems and in particular with the ADAS model are suitably adjusted and displayed. In the I/O subsystem also the errors affecting the signals referring to the control variables are calculated, with reference to the target ones, so that the instantaneous error value can be deduced at any given time.

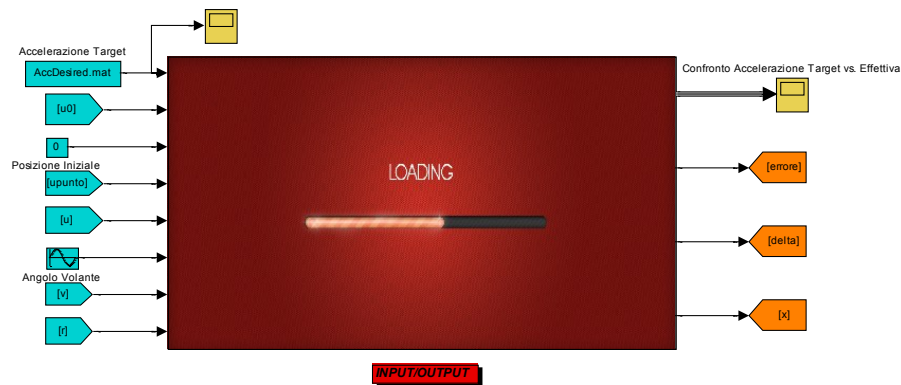


Figure 26 – “Input/Output” Simulink subsystem

Error determination is essential for the vehicle dynamics model, as this variable stands for the input signal controlling the processor units, designated to monitor the vehicle’s behaviour and to make the decisions to adapt its working conditions to the desired ones (e.g. action of throttle and braking system).

In the control unit subsystem the vehicle’s real components and systems behaviour were introduced in order to make the final model as close to reality as possible.

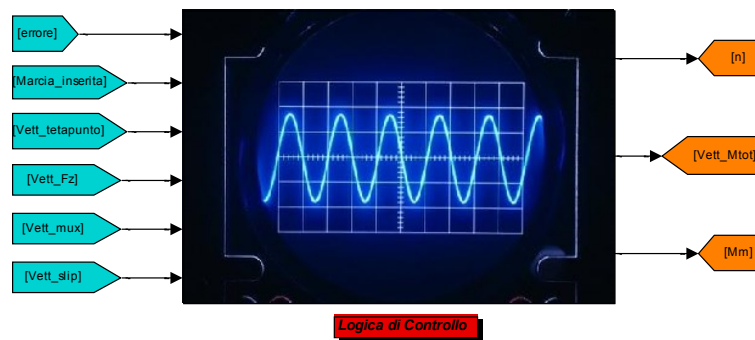


Figure 27 – “Control Unit Simulink subsystem”

In particular, this subsystem allows to take into account the engine’s dynamic behaviour, as well as the behaviour of the braking system, complete with an Antilock Braking System (ABS) unit. The ABS working principle is based on the wheels’ speed detection (through dedicated sensors fitted on board), signal processing and a set of actuators directly operating the hydraulic

braking system. In the ABS block two proportional-integral controllers were used to estimate the magnitude of the action to be performed on the vehicle's systems to let it move according to the ADAS control logics. The input signal for these controllers is the error magnitude, while the output signal is the action, which has to be carried out in order to minimize the error itself.

In the tyre subsystem, an essential part of the whole vehicle model, an evolved Magic Formula set of equations (known also as Pacejka equations) able to simulate wheel behaviour was introduced.



Figure 28 – “Tires” Simulink subsystem

In this subsystem it is possible to determine the tyres' actual grip at any one time and to calculate the longitudinal and lateral interaction forces the tyre is able to perform. The above-mentioned interaction forces are the friction actions the tyres exert on the ground, depending both on the local friction coefficient inside the contact patch, strongly connected to fundamental thermodynamic phenomena observable at the tyre/road interface (Farroni, Giordano, Russo, & Timpone, 2014), (Farroni, Sakhnevych, & Timpone, 2015), both on the vertical force acting on the tyre, hence even depending on the load transfer during vehicle handling. Knowledge of these forces is essential, since they deeply affect the vehicle's handling and road-holding: this means that these forces are intimately connected to vehicle safety, and hints at the importance of this subsystem and of the equations it aims to introduce.

The vehicle dynamic subsystem can be considered as one of the most important parts of the whole model, since it has the role of modelling the

vehicle's behaviour, bringing the ADAS model closer to reality, which is the main aim of the entire model itself.



Figure 29 - “Vehicle Dynamics” Simulink subsystem

Moreover in order to obtain a correct vehicle's behaviour modelling, all the gearbox issues are dealt with.

As it is obvious, being the aim the realistic modelling of a commercial car, the implementation of a gearbox control unit in this model cannot be disregarded: this is a device able to detect the currently engaged gear and to act consequently.



Figure 30 – “Gearbox Control Unit” Simulink subsystem

As a matter of fact, all modern vehicles must be equipped with a gearbox, automatic or manual, in order to let the engine adapt to all kind of roads and situations.

The output signals provided by the vehicle dynamics model, representing input variables for the ADAS model, are the following:

- ‘Applied’: This is a vector containing position, speed and actual acceleration the vehicle is able to develop at any given time, following to the provided input (Desired Acceleration) and considering its real components potentialities, such as the engine, tires, brake system and so on.

- ‘RPM’: This is the simulated rotational speed, calculated using the vehicle dynamics model, the engine has to work at in order to provide the corresponding actual acceleration contained in the ‘Applied’ variable at any given time. It is important to underline this quantity is not the engine rotational speed measured by the OBD system installed on-board, since it is calculated starting from the angular velocities of the driving wheels, being the transmission ratios known variables.

- ‘Throttle’: This is the simulated position the admission regulation device (butterfly valve or injection pump fuel rack) the engine has to work with in order to provide the corresponding actual acceleration contained in the ‘Applied’ variable at any given time.

These outputs correspond to the independent variables of fuel consumption model developed in Chapter 3; in fact these parameters can be used to calculate the instantaneous fuel consumption of the eco-drive system in order to compare it with those observed.

In fact the Fuel Consumption model presented the previous described equations (34) and (46).

4.2 *The implemented model in DIL*

The same architecture implemented in MIL (Model In the Loop) has been implemented in DIL (Driver In the Loop). This has been possible using SCANeR Studio software as virtual environment simulator. The advantages of using a virtual environment are the possibility to have a controlled environment in which the designed ADAS can be tested safely, and also to evaluate the driver-acceptability of a prototyped Human Machine Interface (HMI).

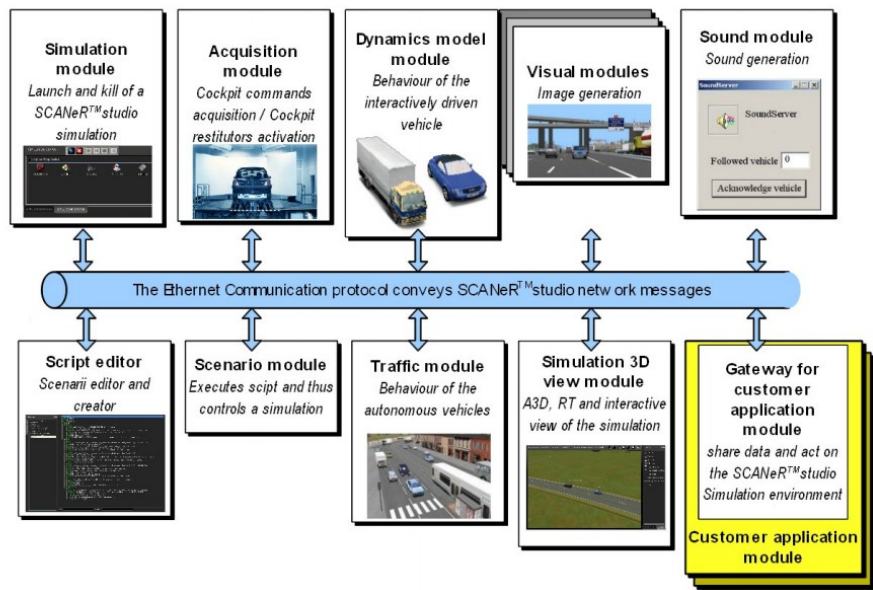


Figure 32 – Simulink/SCANeR communication protocol

A Simulink S-function developed by OKTAL allows interfacing a Simulink model with SCANeR communication protocols. Like every customer made modules, it is based on the SCANeR API (Application Programming Interface). This S- function reads the network and the shared memory, and sends data to the Simulink model. Similarly, it writes on the network and the shared memory data computed by the Simulink model.

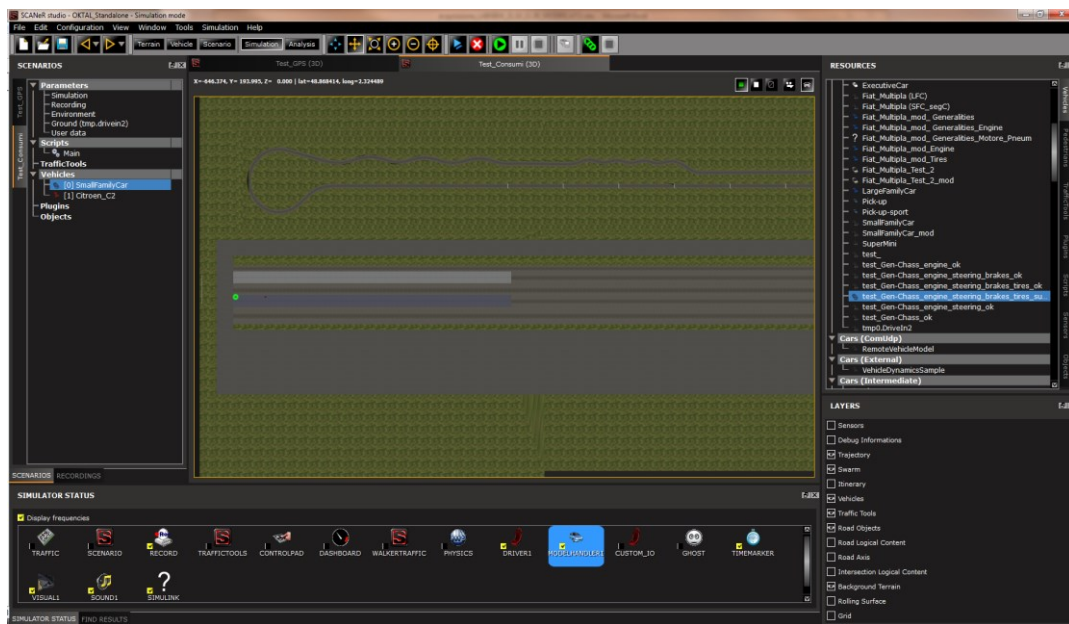


Figure 33 – SCANeR Studio window

In order to have communication between the two software, is essential that both software work in the same folder; in this way it's possible to access at dedicated library for SCANeR-Simulink communication.

For the case under study, the following blocks are used:

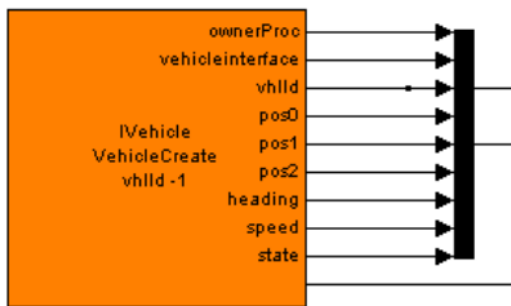
- Controller



which allows Simulink or the generated code to be detected as a SCANeR module (simulation state management and scheduling). Moreover using this block it's possible to set the execution frequency of Simulink model and the network output frequency. The latter must be a sub multiple of the model execution frequency, and it's necessary in order to avoid a network overload when executing a model at high frequency. In this case an execution frequency of 1000 Hz has been set due to vehicle dynamic module that need to run with an high frequency because of differential equations' resolution problems. As network output frequency has been set equal to 20 Hz; due to the fact that the driver receives a feedback according to

minutes. Finally the state and the simulation time are the outputs of controller block.

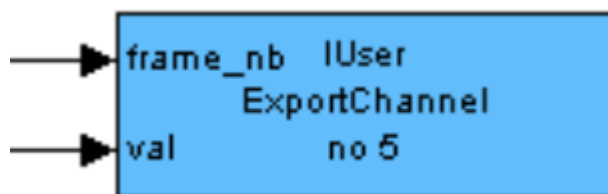
- Network input blocks



it is used to retrieve simulation data in the Simulink mode. As shown in the figure above these blocks have output ports but don't have input ports. Each output port represents a field of the message. When a message is received, all fields are updated at the same time

step. For the developed Simulink model the following input blocks are used: one related to the interactive vehicle in order to have information about the kinematic vehicle parameters (e.g. rpm, throttle, acceleration and speed), another one in order to collect the position of interactive vehicle and finally an input block regarding a radar able to provide kinematic information on the vehicular stream boundary.

- Network output blocks



it is used to send Simulink model outputs to other SCANeR modules. Despite the input blocks, these blocks have input

ports but don't have output ports. Each input port represents a field of the message. A message is always sent with all of its fields. It is important to fill all fields in order not to send incoherent values to other modules. For our purpose, the value calculated in the Simulink model is sent to SCANeR.

Furthermore respect to the model developed for MIL, an additional Simulink block has been inserted, in particular this block implements the fuel consumption model characterized and validated in the previous chapter (Chapter 3). In fact, knowing the observed kinematic and motor conditions it's possible to calculate the fuel consumption of the driver under test in real time

and compare it with the *eco*-drive ones. This is due to the fact that the SCANeR module doesn't provide information about fuel consumption.

Finally in order to activate the exchange of data between Simulink and SCANeR, it was necessary to program the activation of these channels in the main window of SCANeR. Indeed receiving the ratio value of fuel consumption observed and fuel consumption calculated (using the *eco*-drive system) each one-minute; this numerical signal becomes an image signal (Figure 34).

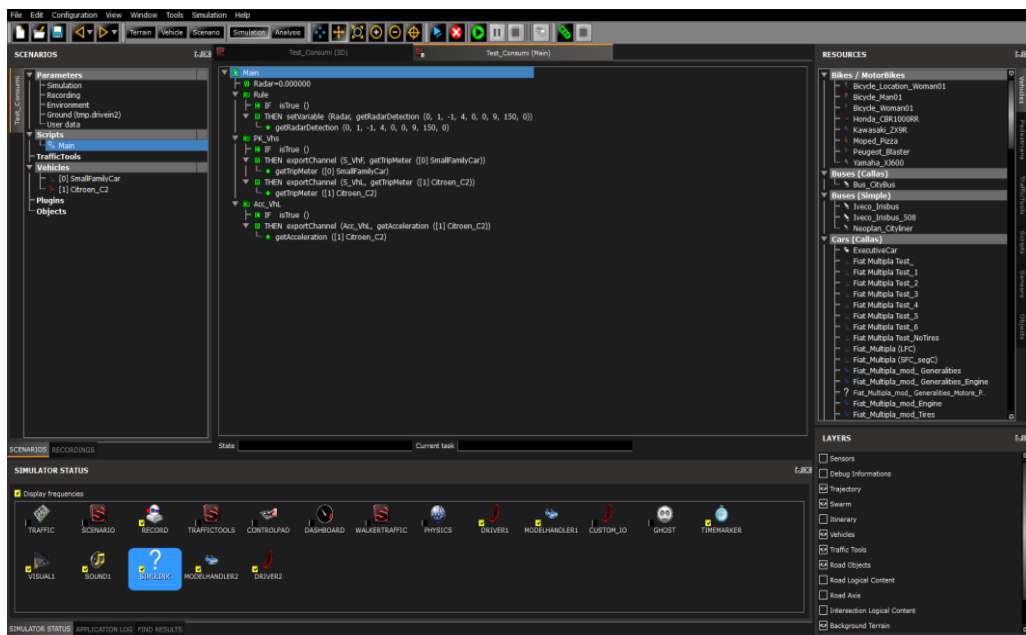


Figure 34 - SCANeR Studio main

In fact on the virtual car dashboard a leaf differently coloured, according to rules of traffic light, appears depending on the value of this ratio. In particular if the ratio is lower than 1.1 the leaf is green. This means that the driver is using fewer diesel than the *eco*-drive system, or almost the same diesel as the *eco*-drive system. If the ratio is strictly less than 1.00, the parameters of the model representing the actual driving behaviour are recorded in order to update the *eco*-ACC. If the ratio is between 1.1 and 1.5 the leaf is yellow, the driver is using more diesel than the *eco*-system in particular s/he is using more than 10-50%. Finally if the driver is using much more diesel than the *eco*-system (ratio value > 1.5) the leaf is red coloured. In order to provide this kind of graphical information an include script file (a sort of s-function in Simulink model) has been invoked in the SCANeR Studio main (Figure 35).



Figure 35 – (on the top) Leaves used for driver feedback, (on the bottom) Screenshot of graphical interface provides to driver

Conclusion

The thesis is oriented to present the work carried out in the set-up of an observation and simulation platform for driving behaviours, aimed to develop and test ADAS solutions. While the work carried out during the thesis has been mainly devoted to the development and testing of the *eco*-driving aspects of the platform, the general tool implements a large variety of functionalities. The platform permits to observe (and understand) driving behaviour with the aim to develop human-like ADAS logics. Logics development is aided by the simulation platform in a trial-and-testing approach (MIL). Also, the platform has a driving simulator that allows investigating the driver acceptability of the ADAS solutions (DIL). As shown during the thesis, the platform integrates several tools: an instrumented vehicle, for the observation of driving behaviour, software tools (e.g. Matlab/Simulink), allowing for a MIL approach, and finally; a driving simulator, to reproduce and simulate the ADAS functions in realistic driving and traffic conditions, as well as to investigate the driver acceptability.

The platform has been employed for the observations of driving behaviour in a large behavioural survey. More than 100 drivers have been observed twice, in the real world and in the virtual environment, where the virtual environment emulates the same context of real world. This also allows a relative validation for the driving simulator performed by Bifulco et al. (2013). For the purpose of the thesis only the data collected using the Instrumented Vehicle are used in order to study the driving behaviour and to specify and validate a fuel consumption model. The fuel consumption model specified is a model able to predict the instantaneous fuel consumptions adopting as independent variables that are very easily estimated by simulation platforms or collected from on-board low cost devices, as well as it is able to identify in real-time the actual value of its parameters. Indeed the same Instrumented Vehicle, equipped with PEMS also, has been used to validate the fuel data collected by the OBD-II port. This is an important step to perform due to the fact that the direct access of data via CAN (Controlled Area Network) is not allowed. In terms of aggregate data, the

results are very satisfactory both in terms of model's specification and correlation between the two measurers (OBD port and PEMS).

Validated the fuel consumption data collected by the OBD-II port and specified/validated the fuel consumption model, the next step is consisted on the development and integration into the testing platform of an *eco*-ACC logic. The implemented *eco*-ACC is a "next-generation" *eco*-drive system; in fact the system is able to perform an *eco*-driving style in the same observed traffic condition. In order to validate the technical applicability of the solution in real-time conditions and to test the driver-acceptability of the solution, MIL (Model In the Loop) and DIL (Driver In the Loop) simulations have been carried out. The first one, based on a desktop-simulation approach (MIL), is employed for the earlier stage of ADAS design. The virtual testing environment is composed by Matlab/Simulink. In order to obtain a realistic traffic conditions and realistic sensors' models, and without an high effort, the implemented system could be implamented in dedicated software as PreScan (strictly speaking it's a sort of Matlab/Simulink with a GUI interface) that allows to implements model for the vehicle sensors and the traffic scenario. The second one, DIL environment, is based on the use of driving simulator, the latter can be adopted to assess the solution in a realistic driving environment, with realistic traffic and under the great variety of different stimuli and conditions. Interestingly, the driving simulator can implement the ADAS logic employing exactly the same Matlab/Simulink blockset employed for the previous part (desktop assessment). The driving simulator software (SCANeR, by OKTAL) allows using different API able to interact with Matlab/Simulink. Moreover it has been possible to implement the HMI (Human Machine Interface) that provides information to driver *under test* about his/her keeping, or not, of an *eco*-driving style. In that case a graphical human machine interface has been developed based on the rules of traffic light in order to be easily understandable (and intuitive) from the drivers. Future works will be oriented to investigate the driver acceptability of the defined HMI.

References

- Ahn, K. (1998). *Microscopic Fuel Consumption and Emission Modeling. Master's Thesis*. Blacksburg, VA, USA: Virginia Polytechnic Institute and State University.
- Ahn, K., & Rakha, H. (2009). A field evaluation case study of the environmental and energy impacts of traffic calming. *Transportation Research Part D*, 14, 411-424.
- Akcelik, R. (1989). Efficiency and drag in the power-based model of fuel consumption. *Transportation Research Part B Methodological*, 23, 376-385.
- Alessandrini, A., Filippi, F., & Ortenzi, F. (2012). Consumption Calculation of Vehicles Using OBD Data. *Centre For Transport and Logistics, University of Rome*.
- Barth, M., An, F., Younglove, T., Scora, G., Levine, C., Ross, M., et al. (2000, April). *Development of a Comprehensive Modal Emissions Model*. Retrieved October 16, 2015, from http://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_w122.PDF
- Bifulco, G. N., Bracco, F., Chiorri, C., Galante, F., Pariota, L., & Russo Spena, M. (2013). Driving simulator: validating the experimental environment for workload research. *Road Safety and Simulation International Conference (RSS2013)*. Rome, Italy.
- Bifulco, G. N., Pariota, L., Galante, F., & Fiorentino, A. (2012). Coupling Instrumented Vehicles and Driving Simulators: opportunities from DRIVE IN2 Project. *15th International IEEE Conference on Intelligent Transportation Systems* (pp. 1815-1820). Anchorage: IEEE.
- Bifulco, G. N., Pariota, L., Simonelli, F., & Di Pace, R. (2011). Real time smoothing of car-following data through sensor fusion techniques. *Procedia - Social and Behavioral Science*, 20, 524-535.
- Bracco, F., & Chiorri, C. (2006). Validazione italiana del NASA-TLX su un campione di motociclisti [Italian validation of the NASA-TLX on a sample of bikers]. *XV Congresso Nazionale dell'Associazione Italiana di Psicologia*. Rovereto, Italy: Sezione di Psicologia Sperimentale.

- Brzezinski, D., & Newell, T. (1998). MOBILE 6 A Revised Model for Estimation of Highway Vehicle Emissions. *Emission Inventory: Living in a Global Environment*, 1, 526–567.
- Carteni, A. (2014). *Urban Sustainable Mobility. Part 1: Rationality in Transport Planning*. Retrieved October 19, 2015, from http://transportproblems.polsl.pl/pl/Archiwum/2014/zeszyt4/2014t9z4_04.pdf
- Chi, Y., Guo, Z., Zheng, Y., & Zhang, X. (2014). Scenarios Analysis of the Energies' Consumption and Carbon Emissions in China Based on a Dynamic CGE Model. *Sustainability*, 6, 487-512.
- Cicero-Fernández, P., Long, J. R., & Winer, A. M. (1997). Effects of Grades and Other Loads on On-Road Emissions of Hydrocarbons and Carbon Monoxide. *Journal of the Air & Waste Management Association*, 47(8), 898-904.
- D'Acierno, L., Montella, B., & Gallo, M. (2011). A fixed-point model and solution algorithms for simulating urban freight distribution in a multimodal context. *Journal of Applied Science*, 11, 647-654.
- Davitashvili, T. (2009). *Mathematical Modeling Pollution from Heavy Traffic in Tbilisi Streets*. Retrieved October 16, 2015, from <http://www.wseas.us/e-library/transactions/environment/2009/29-536.pdf>
- De Luca, S., Di Pace, R., & Marano, M. (2015). Modelling the adoption intention and installation choice of an automotive after-market mild-solar-hybridization kit . *Transportation Research Part C*, 56, 426-445.
- Donateo, T. (2013). CO2 impact of intelligent plugin vehicles. . *WSEAS Transactions on Environment and Development*, 9, 240-252.
- Eggemeier, F. T., Wilson, G. F., Kramer, A. F., & Damos, D. L. (1991). General considerations concerning workload assessment in multi- task environments. *Multiple task performance* , 207-216.
- Farroni, F., Giordano, D., Russo, M., & Timpone, F. (2014). TRT: Thermo racing tyre a physical model to predict the tyre temperature distribution. *Meccanica*, 49(3), 707-723.
- Farroni, F., Russo, M., Russo, R., & Timpone, F. (2014). A physical–analytical model for a real-time local grip estimation of tyre rubber in sliding contact with road asperities. *Proceedings of the Institution of*

Mechanical Engineers, Part D: Journal of Automobile Engineering, 228(8), 955-969.

- Farroni, F., Russo, M., Russo, R., Terzo, M., & Timpone, F. (2013). A combined use of phase plane and handling diagram method to study the influence of tyre and vehicle characteristics on stability. *Vehicle System Dynamics*, 51(8), 1265-1285.
- Farroni, F., Russo, R., Russo, M., Terzo, M., & Timpone, F. (2012). On the influence of anti-roll stiffness on vehicle stability and proposal of an innovative semi-active magnetorheological fluid anti-roll bar. *Raad 2012 Proceedings-21th International Workshop on Robotics in Alpe-Adria-Danube Region*. Naples.
- Farroni, F., Sakhnevych, A., & Timpone, F. (2015). An Evolved Version of Thermo Racing Tyre for Real Time Applications. *World Congress on Engineering*. London.
- Franco García, V. (2014). *Evaluation and improvement of road vehicle pollutant emission factors based on instantaneous emissions data processing*.
- Gałecki, A., & Burzykowski, T. (2013). *Linear Mixed-Effects Models Using R - A Step-by-Step Approach*. Springer.
- Gori, S., La Spada, S., Mannini, L., & Nigro, M. (2014). Emission dynamic meso-simulation model to evaluate traffic strategies in congested urban networks. *IET Intelligent Transport System*, 9, 333-342.
- Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale. *Human mental workload*, 139-183.
- Hui, G., Zhang, Q., Yao, S., & Wang, D. (2007). Evaluation of the International Vehicle Emission (IVE) model with on-road remote sensing measurements. *Journal on Environmental Sciences*, 19, 818-826.
- Jamson, A. H., Hibberd, D. L., & Merat, N. (2015). Interface design considerations for an in-vehicle eco-driving assistance system. *Transportation Research Part C: Emerging Technologies*, 58, 642-656.
- Koupal, L., Cumberworth, M., Michaels, H., Beardsley, M., & Brzezinski, D. (2003). *Design and Implementation of MOVES: EPA's New Generation Mobile Source Emission Model*. Retrieved October 16, 2015, from <http://www3.epa.gov/ttn/chief/conference/ei12/mobile/koupal.pdf>

- Kreft, I., & DeLeeuw, J. (1988). Introducing multilevel modeling. London: Sage Publications Ltd.
- Larsson, H., & Ericsson, E. (2009). The effects of an acceleration advisory tool in vehicles for reduced fuel consumption and emissions. *Transportation Research Part D*, 14, 141-146.
- Lee, M., Park, Y., Jung, K., & Yoo, J. (2011). Estimation of Fuel Consumption using In-Vehicle Parameters. *International Journal of u-and e-Service*, 4(4), 37-46.
- Leung, D., & Williams, D. (2000). Modeling of motor vehicle fuel consumption and emissions using a power-based model. *Urban Air Quality: Measurement, Modelling and Management*, 65, 21-29.
- Matthews, G. D. (1996). Validation of the driver stress inventory and driver coping questionnaire. *International Conference on Traffic and Transport Psychology*. Valencia, Spain.
- Mauro, R., Giuffrè, O., Grana, A., & Chiappone, S. (2014). 11. A statistical Approach for Calibrating a Microsimulation Model for Freeways . *WSEAS Transactions on Environment and Development*, 10, 496-598.
- Ntziachristos, L., Gkatzoflias, D., Kouridis, C., & Samaras, Z. (2009). COPERT: A European road transport emission inventory model. *Information Technologies in Environmental Engineering*, 491-504.
- OKTAL, *SCANeR Studio Manual*.
- Pamanikabud, P. (2009). Impact Assessment of Motorway Traffic Noise Using Visualized Noise Mapping Technique. *WSEAS Transactions on Environment and Development*, 5, 24-33.
- Pariota, L. (2013). Driving behaviour for ADAS: theoretical and experimental analyses - PhD thesis.
- Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2015). nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-122.
- Post, K., Kent, J. H., Tomlin, J., & Carruthers, N. (1984). Fuel consumption and emission modelling by power demand and a comparison with other models. *Transportation Research Part A General*, 18, 191-213.
- R Core Team. (2012).

- Rakha, H. (2015). Transportation Sustainability: What Can Intelligent Transportation Systems Offer? *Engineering & Technology Reference 1.1*.
- Rakha, H., Ahn, K., & Trani, A. (2004). Development of VT-Micro model for estimating hot-stabilized light duty vehicle and truck emissions. *Transportation Research Part D Transportation Environment*, 9, 49-74.
- Rakha, H., Ahn, K., Moran, K., Searens, B., & Van den Bulck, E. (2011). Virginia Tech Comprehensive Power-Based Fuel Consumption Model: Model development and testing. *Transportation Research Part D Transportation Environment*, 16, 492-503.
- Russo Spena, M., Timpone, F., & Farroni, F. (2015). Virtual Testing of Advanced Driving Assistance Systems. *INTERNATIONAL JOURNAL OF MECHANICS*, 9, 300-308.
- Saerens, B., Diehl, M., & Van den Bulck, E. (2010). Automotive Model Predictive Control: Models, Methods and Applications. *Lecture Notes in Control and Information Science*, 402, 119-138.
- Santos Nobre, J., & da Motta Singer, J. (2007). Residual analysis for linear mixed models. *Biometrical Journal*, 49(6), 863-875.
- Sivak, M., & Schoettle, B. (2012). Eco-driving: Strategic, tactical, and operational decisions of the driver that influence vehicle fuel economy. *Transport Policy*, 22, 96-99.
- Stanton, N., Young, M., & McCaulder, B. (1997). Drive-by-wire: the case of driver workload and reclaiming control with adaptive cruise control. *Safety Science*, n.27, 149-159.
- Terracciano, A. M. (2003). Factorial and construct validity of the Italian Positive and Negative Affect Schedule (PANAS). *European Journal of Psychological Assessment*, 19, 131-141.
- Toyota Motor Corporation TMC. (2011). *Toyota FJ Cruiser Features Brake System*. Retrieved November 2014, from <http://www.toyota.com.au/fj-cruiser/features/car-safety/brake-system>
- Van der Voort, M., Dougherty, M., & van Maarseveen, M. (2001). A prototype fuel-efficiency support tool. *Transportation Research Part C*, 9, 279-296.
- Wickens, C. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science*, n. 3(2), 159-177.

- Wong, J. (2001). *Theory of Ground Vehicles*. Hoboken, NJ, USA: John Wiley & Sons.
- Zhang, X., Xie, J., Rao, R., & Liang, Y. (2014). Policy Incentives for the Adoption of Electric Vehicles across Countries. *Sustainability*, 6, 8056-8078.
- Zhou, X., Huang, J., Weifeng, L., & Li, D. (2013). Fuel consumption estimates based on driving pattern recognition. *Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCoM), IEEE International Conference on and IEEE Cyber, Physical and Social Computing*, 496-503.

